unclassified

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION

1a. REPORT SECURITY CLASSIFICATION: unclassified	1b. RESTRIC
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBU: unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S): ISS 9021	5. MONITORING ORGANIZATION REPORT NUMBER(S) ...ed If there
6a. NAME OF PERFORMING ORGANIZATION: Princeton University; 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION: Office of Naval Research
6c. ADDRESS (City, State and ZIP Code): Depart. of Electrical Engineering, Princeton, NJ 08544	7b. ADDRESS (City, State and ZIP Code): 800 N. Quincy, Arlington, VA 22217
8a. NAME OF FUNDING/SPONSORING ORGANIZATION: ONR; 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: N00014-87-k-0054
8c. ADDRESS (City, State and ZIP Code): 800 N. Quincy, Arlington, VA 22217	10. SOURCE OF FUNDING NOS.

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ...NO.
			...00.00.1-R.

11. TITLE (Include Security Classification): The Multiuser Information Theory of CDMA CHannels

12. PERSONAL AUTHOR(S): Sergio Verdu

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Yr., Mo., Day)	15. PAGE COUNT
Final	FROM 11/87 TO 3/31/90	May 30, 1990	109

16. SUPPLEMENTARY NOTATION

17. COSATI CODES		
FIELD	GROUP	SUB-GR.

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number): Spread Spectrum, Code Division Multiple Access, Information Theory, Multiuser communication. CRH/ 5 for a project

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Final Report of ONR Contract N00014-87-k-0054 on the fundamental limits of Code Division Multiple Access communications. It includes the Publications/Patents/Presentations/Honors report; a summary of publications sponsored by the contract and reprints of the key publications sponsored by the contract. Keywords:

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS ☐	unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE NUMBER (Include Area Code) / 22c. OFFICE SYMBOL

) FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE.

OFFICE OF NAVAL RESEARCH

FINAL REPORT

for

JANUARY 1987 - 31 MARCH 1990

or

Contract N00014-87-K-0054

R & T No. 411J010---01

THE MULTIUSER INFORMATION THEORY OF CODE-DIVISION MULTIPLE-ACCESS CHANNELS

Sergio Verdú, Principal Investigator
Department of Electrical Engineering
Princeton University
Princeton, New Jersey 08544
Phone: (609) 258-5315
Email: Verdu@princeton.edu

OFFICE OF NAVAL RESEARCH
PUBLICATIONS/PATENTS/PRESENTATIONS/HONORS REPORT
for
1 JANUARY 1987 - 31 MARCH 1990

R&T Number: 411J010---01

Contract/Grant Number: N00014-87-K-0054

Contract/Grant Title: The Multiuser Information Theory of Code-Division Multiple-Access Channels

Principal Investigator: Sergio Verdú

Mailing Address: Department of Electrical Engineering
Princeton University
Princeton, NJ 08544

Phone Number: (with Area Code) (609) 258-5315

E-Mail Address: verdu@princeton.edu

a. Number of Papers Submitted to Refereed Journal but not yet published: 4

b. Number of Papers Published in Refereed Journals: 6
list attached)

c. Number of Books or Chapters Submitted but not yet Published: 1

d. Number of Books or Chapters Published: 0
list attached)

e. Number of Printed Technical Reports & Non-Refereed Papers: 18
list attached)

f. Number of Patents Filed: 0

g. Number of Patents Granted: 0
list attached)

h. Number of Invited Presentations at Workshops or Prof. Society Meetings: 5

i. Number of Presentations at Workshops or Prof. Society Meetings: 18

j. Honors/Awards/Prizes for Contract/Grant Employees:
list attached, this might include Scientific Soc Awards/Offices, 16
Promotions, Faculty Awards/Offices etc)

k. Total number of Graduate Students and Post-Docs Supported at least 25% this year on this contract/grant:

Grad Students 3 and Post-Docs _____.

How many of each are females or minorities? These 6 numbers are for ONR's EEO/Minority Reports; minorities include Blacks, Aleuts AmIndians, etc and those of Hispanic or Asian extraction or nationality. The Asians are singled out to facilitate the varying report semantics re "economically disadvantaged").

Grad Student Female _____

Grad Student Minority _____

Grad Stu Asian Ext 0

Post-Doc Female _____

Post-Doc Minority _____

Post-Doc Asian Ext. _____

Title:	The Multiuser Information Theory of CDMA Channels
Institution:	Department of Electrical Engineering School of Enginnering / Applied Science Princeton University Princeton, NJ 08544
Principal Investigator:	Sergio Verdú, Associate Professor
Social Security No.:	[redacted]
Office Telephone:	(609)-258-5315
Starting Date:	January 1, 1987
Closing Date:	March 31, 1990

SUMMARY OF WORK SPONSORED BY CONTRACT N00014-87-K-0054

A. Refereed Journal Papers not yet published

D. Brady, S. Verdú, "A Semi-Classical Analysis of Optical CDMA," *IEEE Trans. Communications.* vol. COM-39, to appear.

S. Verdú. "On Channel Capacity per Unit Cost," *IEEE Trans. Information Theory,* vol. IT-36, Sep. 1990, to appear.

T. Cheng and S. Verdú, "Capacity of RMS Bandlimited Gaussian Channels," *IEEE Trans. Information Theory,* to appear.

T. Cheng, S. Verdú, "The effect of symbol-asynchronism on multiaccess capacity," *IEEE Trans. Information Theory,* under review.

B. Papers Published in Refereed Journals

S. Ghez, S. Verdú, S. Schwartz, "Stability Properties of Slotted-Aloha with Multipacket Reception Capability," *IEEE Trans. Automatic Control,* vol AC-33,

no. 7, p. 640-649, July 1988.

D. Brady, S. Verdú, "Performance Analysis of an Asymptotically Quantum-Limited Optical DPSK Receiver," *IEEE Trans. Communications,* vol. COM-37, no.1, p. 46-51, Jan. 1989.

S. Verdú, "Multiple-Access Channels with Memory with and without Frame-Synchronism," *IEEE Trans. Information Theory,* vol. IT-35, no. 3, p. 605-619, May 1989.

S. Verdú, "The Capacity Region of the Symbol-Asynchronous Gaussian Multiple-Access Channel," *IEEE Trans. Information Theory,* vol. IT-35, no. 4, p. 733-751, July 1989.

A. Ephremides, S. Verdú, "Control and Optimization Methods in Communication Network Problems." *IEEE Trans. Automatic Control,* vol. AC-34, no. 9, p. 930-942. Sep. 1989.

S. Ghez, S. Verdú, S. Schwartz, "Optimal Decentralized Control in the Multipacket Channel," *IEEE Trans. Automatic Control,* vol. AC-34, no. 11, p. 1153-1163, Nov. 1989.

C. Book Chapters not yet published

S. Verdú, "Multiuser Detection," *Advances in Statistical Signal Processing, vol. 2: Signal Detection* H. V. Poor and J. B. Thomas, Editors, JAI Press: New York, 1990

E. Conference Publications

S. Ghez, S. Verdú and S.C. Schwartz, "Stability of Multipacket Aloha," *Proc. 21st. Conf. Information Sciences and Systems,* The Johns Hopkins University, Baltimore, MD, p. 262-267, Mar. 1987.

S. Verdú, "Fundamental Limits of Code Division Multiple Access Channels," Invited paper, *1987 IEEE Information Theory Workshop,* Bellagio, Italy, June 1987.

S. Ghez, S. Verdú and S. Schwartz, "Decentralized Control Algorithms for Multipacket Aloha," *Proc. Twenty-fifth Allerton Conf. on Communication, Control and Computing,* p. 1089-1098, Oct. 1987.

S. Verdú, "Capacity of Frame-Asynchronous Multiple-Access with Memory," *Abs. 1988 IEEE International Symposium on Information Theory,* Kobe, Japan, p. 81, June 1988.

D. Brady, S. Verdú, "Probability of Error of an Optical DPSK Detector under Brownian Laser Phase Noise," *Abs. 1988 IEEE International Symposium on Information Theory,* Kobe, Japan, p. 198, June 1988.

S. Ghez. S. Verdú, "Ergodicity of Controlled and Uncontrolled Multipacket Aloha." *Proc. 1988 Beijing International Workshop on Information Theory,* Beijing, China, p. BII-1.1-4, July 1988.

S. Verdú. "The Multiuser Information Theory of Code-Division Multiple-Access Channels." *ONR Workshop on Computer Networks,* Clemson, SC, Aug. 1988.

S. Verdú, "Capacity of Asynchronous Code-Division Multiple-Access Channels," *Fourth Annual Cornell Summer Workshop on Systems, Control and Communications,* Cornell University, Ithaca, NY, Aug. 1988.

R. S. Cheng and S. Verdú, "Optimal Signal Design for Bandlimited PAM Synchronous Multiple-Access Channels," *Proc. 23rd. Conf. Information Sciences and Systems,* The Johns Hopkins University, Baltimore, MD, p. 321-326, Mar. 1989.

D. Brady and S.Verdú, "An Exact Analysis of the Optical CDMA Noncoherent Receiver," *Proc. 23rd. Conf. Information Sciences and Systems,* The Johns Hopkins University, Baltimore, MD, p. 327-332 Mar. 1989.

S. Verdú, "Completely Asynchronous Multiple-Access Channels," Invited paper, *1989 Workshop on Information Theory,* Cornell University, Ithaca, NY, Jun. 1989

D. Brady and S. Verdú, "Error Probability of CDMA in the Poisson Channel," Invited paper, *1989 Workshop on Information Theory,* Cornell University,

Ithaca, NY, Jun. 1989

S. Verdú, "Capacity of Completely Asynchronous PAM Gaussian Multiple-Access Channels," *Proc. Fourth Joint Swedish-Soviet International Workshop on Information Theory,* p. 370-374, Gotland, Sweden, Aug. 1989

R. Cheng and S. Verdú, "Total Capacity of the RMS Bandlimited K-user PAM Synchronous Channel," *Twenty-Seventh Allerton Conf. on Communication, Control and Computing,* Sep. 27-29, 1989

A. Ephremides and S. Verdú, "Control Issues in Communication Networks," *1989 IEEE Decision and Control Conference,* Tampa, Fl., Dec. 1989

R. Cheng and S. Verdú, "Capacity of RMS Bandlimited Gaussian Multiple-Access Channels," *1990 IEEE International Symposium on Information Theory* Jan. 14-19, 1990

S. Verdú, "The First Forty Years of the Shannon Theory," *Conference Record IEEE TAB Region 10 Tour,* Piscataway, NJ: IEEE Press, 1990. Also, reprinted in *IEEE Information Theory Newsletter,* June 1990.

D. Brady and S. Verdú, "Semi-classical analysis of Optical CDMA receivers," *IEEE/LEOS Optical Multiple Access Networks,* Invited paper, Monterey, CA, Jul. 1990

J. Honors/Awards/Prizes for Contract Employees

S. Verdú, Principal Invesigator

A. Rheinstein Award 1987, Outstanding Junior Faculty Member,
School of Engineering and Applied Science, Princeton University.

NSF Presidential Young Investigator Award, 1988

IEEE Senior Member, 1988

'romoted to Associate Professor, Princeton University, 1989

EEE Information Theory Society Board of Governors, Elected 1989

Engineering Council Award for Excellence in Teaching, 1989
School of Engineering and Applied Science, Princeton University.

Appointed Visiting Fellow, Institute of Advanced Studies,
Australian National University, Canberra, Australia, Sep-Oct 1989

Associate Editor,
IEEE Transactions on Automatic Control, 1988-1990.

Co-Director,
22nd Conference on Information Sciences and Systems, Princeton, NJ, March 1988.

Delegate of the IEEE Information Theory Society, IEEE Region 10 Tour, Oct. 1989,

Member,
Awards Committee, IEEE Technical Activities Board, 1990

Co-Chairman,
1990 IEEE Information Theory Workshop, Veldhoven, The Netherlands, June 1990

Program Chairman,
992 IEEE Conf. Decision and Control, Tucson, Arizona, Dec. 1992

. Ghez, Research Assistant

Wallace Memorial Fellowship in Engineering,
Graduate School, Princeton University, 1987-88

. Cheng, Research Assistant

Meitec Corporation Junior Fellowship Award, 1989

George Van Ness Lothrop Fellowship,
Graduate School, Princeton University, 1990-91

Stability Properties of Slotted Aloha with Multipacket Reception Capability

Sylvie Ghez
Sergio Verdú
Stuart C. Schwartz

Reprinted from
IEEE TRANSACTIONS ON AUTOMATIC CONTROL
Vol. 33, No. 7, July 1988

Stability Properties of Slotted Aloha with Multipacket Reception Capability

SYLVIE GHEZ, STUDENT MEMBER, IEEE, SERGIO VERDÚ, MEMBER, IEEE, AND STUART C. SCHWARTZ, SENIOR MEMBER, IEEE

Abstract—The stability of the Aloha random access algorithm in an infinite-user slotted channel with multipacket reception capability is considered. This channel is a generalization of the usual collision channel, in that it allows the correct reception of one or more packets involved in a collision. The number of successfully received packets in each slot is modeled as a random variable which depends exclusively on the number of simultaneous attempted transmissions. This general model includes as special cases channels with capture, noise, and code division multiplexing. It is shown by means of drift analysis that the channel backlog Markov chain is ergodic if the packet arrival rate is less than the expected number of packets successfully received in a collision of n as n goes to infinity. Finally, the properties of the backlog in the nonergodicity region are examined.

I. Introduction

One of the main problems in random access communications is the determination of the maximum stable throughput. In particular, an important result is that the Aloha protocol is unstable [1]–[3] in an infinite-user slotted collision channel where a transmission is successful only if no other users attempt transmissions simultaneously. Several strategies have been designed to stabilize this channel, such as collision resolution algorithms (see [4], for example) where transmissions are deferred until the current conflict is solved, and more recently, Aloha-type strategies using decentralized control, where the retransmission probability is updated according to previous channel outcomes. It has been shown [5]–[7] that the maximum stable throughput achievable by such Aloha-type strategies with decentralized control is e^{-1}.

However, the collision channel model does not hold in many important practical multiuser communication systems [8]–[21] because simultaneous transmission of several packets does not necessarily result in the destruction of all the transmitted information. For instance, the capture phenomenon is common in local area radio networks [12]–[15]; if the power of one of the received packets is sufficiently large compared to the power of the other packets involved in a collision, then the strongest packet can be correctly decoded, while the other packets are lost. Other examples are multiple-access channels where several users transmit simultaneously in the same frequency band, and a multiuser detector demodulates the information transmitted by all active users (e.g., [8]–[11]). Although those systems do not necessarily require a random access protocol, it is sometimes useful to exercise some flow control through such a protocol so as to limit the maximum number of simultaneous transmitters, in order to bound the multiuser receiver complexity and guarantee lower bit-error rates.

Manuscript received July 23, 1987; revised January 8, 1988. Paper recommended by Past Associate Editor A. Ephremides. This work was supported in part by the Office of Naval Research under Contract N00014-87-k-0054 and by the Army Research Office under Contract DAAL03-87-k-0062.

The authors are with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544.

IEEE Log Number 8821359.

Previous studies of some of the aforementioned systems [9], [12]–[18] where some of the packets involved in a collision may be correctly received have shown that the performances are noticeably improved with respect to slotted Aloha. However, even in those special cases, no precise stability result is available, either because finite population networks with no buffer space were considered, or because the Poisson approximation of channel traffic was used for infinite population networks. In [19] (see also [20]), upper and lower bounds are derived for the capacity of a multiple access channel where all packets are correctly received if the collision size does not exceed a fixed threshold and otherwise all packets are destroyed.

In this paper, we consider a generalization of the collision channel, where the receiver can demodulate several packets simultaneously. It is assumed that the number of correctly demodulated packets is a random variable, which, given the number of packets simultaneously transmitted, is independent of the backlog and of the number of previous retransmission attempts. This random variable can take any integer value between zero and the collision size. Thus, the channel is described by a matrix of conditional probabilities (ϵ_{nk}) where ϵ_{nk} is the probability that k packets are correctly demodulated given that there were n simultaneous transmissions. We analyze the usual Aloha algorithm with the multipacket reception capability just described. Users are synchronized so that transmissions take place within one slot, and at the end of each slot, stations that did transmit a packet learn whether or not their transmission was successful. Unsuccessful or backlogged packets are retransmitted in each subsequent slot with probability p, $0 < p \leq 1$. It turns out that multipacket reception capability can stabilize Aloha. Our main result states that the maximum stable throughput is equal to the limit of the average number of packets correctly received in collisions of size n when n goes to infinity. To show this, we model the channel backlog as a Markov chain, and then study its properties by using some simple drift analysis techniques.

The last part of this paper is a study of the properties of the backlog in the nonergodicity region. Unlike the backlog Markov chain for slotted Aloha which is always transient [1], the backlog for our model does in general have a null recurrence region of positive length, which depends on the matrix (ϵ_{nk}) and on the retransmission probability p. However, transience in the nonergodicity region can be ensured for a large class of systems, and in particular for channels where the number of successful simultaneous transmissions is bounded.

II. Multipacket Reception Model

Let A_k be the number of new packets arriving during time slot k. Assume that $(A_k)_{k\geq 0}$ are i.i.d. random variables with probability distribution:

$$P[A_k = n] = \lambda_n \qquad (n \geq 0)$$

such that the mean arrival rate $\lambda = \sum_{n=1}^{\infty} n\lambda_n$ is finite. New packets are transmitted with probability one at the beginning of the first slot following their arrival.

Given that n packets are being transmitted in one slot, we define

0018-9286/88/0700-0640$01.00

for $n \geq 1$, $0 \leq k \leq n$,

$$\epsilon_{nk} = P[k \text{ packets are correctly received} | n \text{ are transmitted}].$$

The multipacket reception properties of the channel are summarized by the stochastic matrix

$$E = \begin{bmatrix} \epsilon_{10} & \epsilon_{11} & & \\ \epsilon_{20} & \epsilon_{21} & \epsilon_{22} & 0 \\ \cdot & \cdot & & \\ \epsilon_{n0} & \epsilon_{n1} & & \epsilon_{nn} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \end{bmatrix}$$

which we refer to as the *reception matrix* of the channel. For instance, the reception matrix for the usual collision channel is

$$\begin{bmatrix} 0 & 1 & \\ 1 & 0 & 0 \\ 1 & 0 & \\ 1 & \cdot & \\ \cdot & \cdot & \\ \cdot & \cdot & \end{bmatrix}$$

while for a system with capture it has the form

$$\begin{bmatrix} 0 & 1 & \\ 1-x_2 & x_2 & 0 \\ \cdot & \cdot & \\ 1-x_n & x_n & \\ \cdot & \cdot & \\ \cdot & \cdot & \end{bmatrix}$$

where x_n is the probability of capture given that the collision size is n. The model studied in [19], [20] can be described by a reception matrix of the form

$$\begin{bmatrix} 0 & 1 & & & \\ 0 & 0 & 1 & & 0 \\ \cdot & \cdot & & \cdot & \\ 0 & 0 & & & 1 \\ 1 & 0 & & & \\ 1 & 0 & & & 0 \\ \cdot & \cdot & & & \end{bmatrix}.$$

Note that by letting $\epsilon_{10} \neq 0$ our model allows not only collisions but also background noise to be a source of errors.

Denote by X_n the number of backlogged packets in the system at the beginning of slot n. The discrete-time process $(X_n)_{n \geq 0}$ is easily seen to be a homogeneous Markov chain. We define the system to be stable if $(X_n)_{n \geq 0}$ is ergodic and unstable otherwise. The average number of packets correctly received in collisions of size n is denoted by $C_n = \sum_{k=1}^{n} k\epsilon_{nk}$. We can now state the main result.

Theorem 1: If C_n has a limit $C = \lim_{n \to \infty} C_n$, then[1] the system is stable for all arrival distributions such that $\lambda < C$ and is unstable for $\lambda > C$. This also holds if C is infinite: if $\lim_{n \to \infty} C_n = +\infty$, then the system is always stable.

The proof is given in Section III. In the remainder of this section, we use Theorem 1 to analyze several simple random access channels that fall within the scope of the multipacket reception channel.

1) Mobile Users with Pairwise Transmissions: Consider an infinite number of transmitters $T_1, T_2, \cdots$, and an infinite number of receivers $R_1, R_2, \cdots$, whose positions in the plane are i.i.d. random variables. Suppose that transmissions are pairwise

[1]This result holds under the assumption that the Markov chain of the number of backlogged packets is irreducible and aperiodic (for details and sufficient conditions, see Section III).

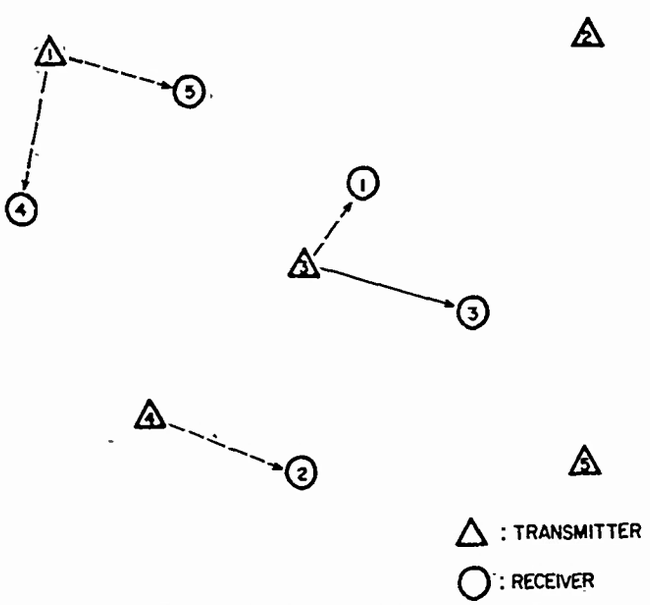

Fig. 1. Pairwise transmissions with only one success (3-3).

in the sense that transmitter T_n sends packets only to receiver R_n, and R_n is only interested in the packets sent by T_n (see Fig. 1). Assume also that each receiver can only detect correctly the packet sent by the closest transmitter (in particular, this is the case if there is perfect capture, see Example 3 below). The successes of transmissions occurring at the same time are independent, so that for $n \geq 2$

$$\epsilon_{nk} = \binom{n}{k} p(n)^k (1-p(n))^{n-k}$$

where $p(n)$ is the probability that any given transmitter is successful in a collision of size n, which is equal to $1/n$ if we assume that all locations are memoryless, i.e., independent from slot to slot. It follows that

$$C_n = np(n) = 1$$

and the maximum throughput is 1. More generally, if because of channel noise, the message of the closest transmitter is received correctly with probability α (in other words $\epsilon_{11} = \alpha$), then the throughput is equal to α. The assumption that the locations of the stations are memoryless is equivalent to assuming that they move infinitely fast. If this simplifying assumption is dropped, then the number of successes depends not only on the current number of retransmissions, but also on the previous history of retransmissions, and thus the problem is no longer encompassed by our multipacket reception model. In Fig. 2, the result of a simulation shows that for moderate speeds, the actual throughput is well approximated by the foregoing analysis.

2) Frequency Hopping Random Access Channel: Consider a finite population of N users transmitting by frequency hopping, as in [11], [22]. For each packet he wants to transmit, a user selects with equal probability one frequency in a fixed set of q frequencies. A packet is correctly received iff no other packet is transmitted on the same frequency during the same slot. We compute $(\epsilon_{Nk})_{1 \leq k \leq N}$, and $C = \lim_{N \to \infty} C_N$. If the users have infinite buffer space, then C can be taken as a good approximation for large N of the maximum stable throughput of the system, which is unknown. If the users have no buffer space, as is often assumed, the backlog Markov chain is always ergodic, but even then, one should expect reasonable delays in large population problems only for arrival rates below C. The computation of the reception matrix of this channel is a simple combinatorial problem of random assignment of objects to cells (e.g., see [23, App. A]).

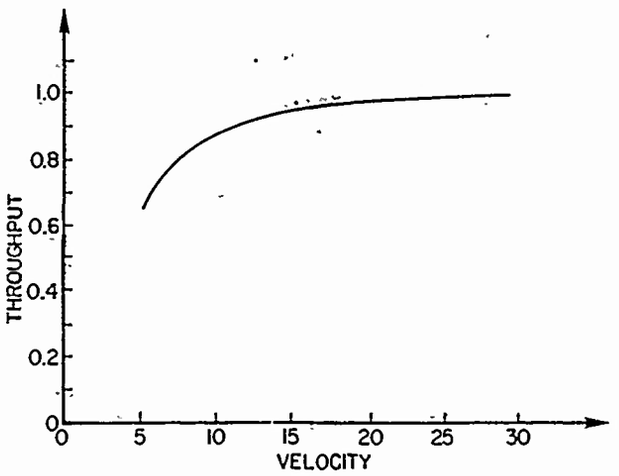

Fig. 2. Throughput as a function of velocity for mobile users with pairwise transmissions. Stations moving in a square region; velocity units: percentage of square side traveled in one slot. Retransmission probability set to 0.1.

Denote by $T_1, T_2, \cdots, T_N$ the users, all involved in the collision, and also denote by S the set of users whose packets are correctly received. Two cases need to be considered.

a) $2 \le N \le q$: We have, for $1 \le j \le N$

$$\epsilon_{Nj}=\binom{N}{j} P[S=\{T_1, T_2, \cdots, T_j\}] \tag{1}$$

and the following decomposition:

$$P[\{T_1, T_2, \cdots, T_j\} \subseteq S]=P[S=\{T_1, T_2, \cdots, T_j\}] + P\left[\bigcup_{k=j+1}^{N} \{\{T_1, \cdots, T_j, T_k\} \subseteq S\}\right]$$

easily yields the desired expression

$$P[S=\{T_1, T_2, \cdots, T_j\}]=\sum_{k=0}^{N-j} (-1)^k \cdot \binom{N-j}{k} P[\{T_1, T_2, \cdots, T_{k+j}\} \subseteq S] \tag{2}$$

where only one term is left to compute

$$P[\{T_1, T_2, \cdots, T_{k+j}\} \subseteq S] = \frac{q(q-1)\cdots(q-j-k+1)(q-j-k)^{N-j-k}}{q^N} \tag{3}$$

for $1 \le j \le N$, $0 \le k \le N-j$. Putting (1), (2), and (3) together gives the result

$$\epsilon_{Nj}=\binom{N}{j} \sum_{k=0}^{N-j} (-1)^k \binom{N-j}{k} \cdot \frac{q(q-1)\cdots(q-j-k+1)(q-j-k)^{N-j-k}}{q^N} \tag{4}$$

for $1 \le j \le N$. Notice in particular that $\epsilon_{N,N-1} = 0$. Let us now compute the average number of packets correctly received in collisions of size N, $C_N = \sum_{j=1}^{N} j\epsilon_{Nj}$. By using (4) and summing at $j + k$ constant, we get

$$q^N C_N=\sum_{i=1}^{N} q(q-1)\cdots(q-i+1)(q-i)^{N-i} \cdot \sum_{n=0}^{i-1} (-1)^n \frac{N!}{n!(i-n-1)!(N-i)!}$$

which can be simplified as

$$q^N C_N=\sum_{i=1}^{N} \frac{N!}{(i-1)!(N-i)!} \cdot q(q-1)\cdots(q-i+1)(q-i)^{N-i}(1-1)^{i-1}$$

to get the final result.

$$C_N=N\left(1-\frac{1}{q}\right)^{N-1}.$$

b) $N > q$: In this case, there can be at best $q-1$ successes in a collision of size N. The same method applies to get the following probabilities:

$$\epsilon_{Nj}=\binom{N}{j} \sum_{k=0}^{q-j-1} \binom{N-j}{k} (-1)^k \cdot \frac{q(q-1)\cdots(q-j-k+1)(q-j-k)^{N-j-k}}{q^N} \qquad (1 \le j \le q-1)$$

$$\epsilon_{Nj}=0 \qquad (q \le j \le N)$$

resulting in the same expected number of successes as before

$$C_N=N\left(1-\frac{1}{q}\right)^{N-1}.$$

Now we let the population size N go to infinity and we apply our result. If we let N grow to infinity while keeping q constant, we have $\lim_{N\to\infty} C_N = 0$, so the system is always unstable. On the other hand, if we let N go to infinity while keeping q equal to a fixed percentage of the population size, i.e., N/q constant, then $\lim_{N\to\infty} C_N = +\infty$, and the system is always stable. It is easily shown that to get a finite maximum stable throughput, q has to grow as $N/\ln N$.

3) Mobile Radio Network with Capture: Consider an infinite number of users independently and uniformly distributed in a circle of radius R, whose positions are independent from slot to slot. Users transmit packets to a common receiver located at the center of the network. Denote by P_1 and P_2 the received powers of the strongest and the next to strongest packets involved in a collision. Assume, as in [12]-[14], that the strongest packet is correctly received iff $P_1/P_2 > K$ (K being a system dependent constant), and that all the other packets involved in the collision are not received successfully. Assume, moreover, that the received power of a packet only depends on the distance r between the sender and the receiver

$$P=\frac{\text{constant}}{r^\alpha} \qquad (\alpha \ge 2).$$

Then there will be capture iff

$$r_2 > \beta r_1$$

where $\beta = K^{1/\alpha}$ is the capture parameter, and r_1, r_2 are the distances of the closest and the next to closest senders from the receiver.

Denote by D the distance between a given user and the receiver. It is easily shown that the pdf of D is given by

$$p_D(r)=2\frac{r}{R^2} \qquad (0\le r\le R).$$

Given N users, denote by U_N the closest from the receiver, and by D_N its distance from the receiver. Computing the cdf of D_N and taking its derivative, we obtain

$$p_{D_N}(r)=2N\frac{r}{R^2}\left[1-\left(\frac{r}{R}\right)^2\right]^{N-1} \qquad (0\le r\le R). \tag{5}$$

Given $D_N = r$, the other $N-1$ users are uniformly distributed in the annular region (r, R). So if N users collide and $D_N = r$, U_N will be correctly received iff all the other users are in the annular region $(\beta r, R)$, which is empty if $\beta r > R$. Therefore, if we denote by

$$P_N(r)=P[\text{capture}|N \text{ collide}, D_N=r] \qquad (N\ge 2)$$

we have

$$P_N(r)=\begin{cases}\left[\dfrac{R^2-\beta^2r^2}{R^2-r^2}\right]^{N-1} & \text{if } r\le\dfrac{R}{\beta}\\ 0 & \text{if } r\ge\dfrac{R}{\beta}\end{cases}. \tag{6}$$

Thus, the probability of capture in a collision of N ($N \ge 2$) is

$$\epsilon_{N1}=\int_0^{R/\beta} P_N(r)p_{D_N}(r)\,dr.$$

Using (5) and (6), and with the change of variable $x = \beta/R$, this is easily computed

$$\epsilon_{N1}=\int_0^{1/\beta} 2Nx(1-\beta^2x^2)^{N-1}\,dx=\frac{1}{\beta^2}.$$

It follows that $C = 1/\beta^2$ is the maximum stable throughput. Notice, in particular, that for $\beta = 1$ (perfect capture), we have $C = 1$ and for $\beta \to \infty$ (no capture), we have $C \to 0$.

Under certain conditions, the performances of Aloha in the multipacket channel can be improved by varying the retransmission probability as a function of the channel history, and a maximum stable throughput of $\sup_{x\ge 0} e^{-x}\Sigma_{n=1}^{\infty} C_n/n!x^n$ can be reached (see [31]).

III. Ergodicity Region

The number of backlogged packets in the system at time n, $(X_n)_{n\ge 0}$, is a homogeneous Markov chain whose one-step transition probability matrix can be computed as a function of p, $(\lambda_k)_{k\ge 0}$, and E. Denoting by $B_i(j)$ the probability of having j retransmissions out of i backlogged packets

$$B_i(j)=\binom{i}{j}p^j(1-p)^{i-j} \tag{7}$$

we get

$$P_{00}=\lambda_0+\sum_{n=1}^{\infty}\lambda_n\epsilon_{nn}$$

$$P_{0k}=\sum_{n=0}^{\infty}\lambda_{k+n}\epsilon_{k+n,n} \qquad (k\ge 1)$$

and for $i \ge 1$

$$P_{i,i-k}=\sum_{n=0}^{\infty}\lambda_n\sum_{j=k}^{i}B_i(j)\epsilon_{n+j,n+k} \qquad (1\le k\le i)$$

$$P_{ii}=\lambda_0\left[B_i(0)+\sum_{j=1}^{i}B_i(j)\epsilon_{j0}\right]+\sum_{n=1}^{\infty}\lambda_n\sum_{j=0}^{i}B_i(j)\epsilon_{n+j,n}$$

$$P_{i,i+k}=\sum_{n=0}^{\infty}\lambda_{k+n}\sum_{j=0}^{i}B_i(j)\epsilon_{j+k+n,n} \qquad (k\ge 1). \tag{8}$$

Sufficient conditions for $(X_n)_{n\ge 0}$ to be irreducible and aperiodic are as follows:

- if $0 < p < 1$:

$$\lambda_0\ne 0 \tag{9a}$$

$$\lambda_0+\sum_{n=1}^{\infty}\lambda_n\epsilon_{nn}<1 \tag{9b}$$

$$\epsilon_{10}\ne 1 \tag{9c}$$

- if $p = 1$:

$$\lambda_0\ne 0 \tag{9a}$$

$$\lambda_0+\sum_{n=1}^{\infty}\lambda_n\epsilon_{nn}<1 \tag{9b}$$

$$\text{for all } i\ge 1,\ \epsilon_{i0}\ne 1. \tag{9d}$$

These are only sufficient conditions, but they hold for almost all nontrivial systems. For example, if (9b) does not hold, then zero is an absorbing state, since the left-hand side of (9b) is equal to P_{00}. Also, (9c) simply means that the successful reception of a single packet in the absence of other active users is possible. Assume, for instance, that $0 < p < 1$ and that the arrivals are Poisson distributed. Then we only have to assume (9c), and (9b) is true unless there is perfect reception, that is $\epsilon_{nn} = 1$ for all $n \ge 1$, in which case the system would of course always be stable. The case $p = 1$ gives rise to a number of pathological situations, hence, the much stronger condition (9d). It generally turns out that either (9d) is not necessary or the stability region of the system is obvious. For instance, it is clear from the transition probabilities that slotted Aloha with $p = 1$ is always unstable. In any case, it is assumed in what follows that $(X_n)_{n\ge 0}$ is irreducible and aperiodic.

Proof of Theorem 1: The proof is based on drift analysis. Recall that in general, the drift at state i ($i \ge 0$) is defined by

$$d_i=E[X_{t+1}-X_t|X_t=i].$$

If we denote by Σ_t the number of successful transmissions in slot t, we have

$$X_{t+1}-X_t=A_t-\Sigma_t$$

and therefore

$$d_i=\lambda-E[\Sigma_t|X_t=i]. \tag{10}$$

Now if R_t is the number of retransmissions in slot t, we get

$$P[\Sigma_t=k|X_t=i,\ A_t=n,\ R_t=j]=\epsilon_{n+j,k}$$

for $0 \le j \le i$, $0 \le k \le n + j$ and with the convention that $\epsilon_{00} =$

$C_0 = 0$. Thus,

$$E[\Sigma_t | X_t = i, A_t = n, R_t = j] = C_{n+j}$$

and

$$E[\Sigma_t | X_t = i] = \sum_{n=0}^{\infty} \lambda_n \sum_{j=0}^{i} B_i(j) C_{n+j}. \tag{11}$$

The value of the drifts for our model follows from (10) and (11)

$$d_i = \lambda - \sum_{n=0}^{\infty} \lambda_n \sum_{j=0}^{i} B_i(j) C_{n+j}. \tag{12}$$

The idea of the proof is to compute $\lim_{i\to\infty} d_i$ which will turn out to be a very simple expression, and then apply the results of [3] and [24] to determine the ergodicity region of $(X_n)_{n\geq 0}$. Let us first recall the two results that will be used in the sequel.

Lemma A (Pakes [24]): Let $(X_n)_{n\geq 0}$ be an irreducible and aperiodic Markov chain having as state space the nonnegative integers, denote by (P_{ij}) its transition probability matrix, and by d_i its drift at state i. Then if for all i $|d_i| < \infty$, and if $\limsup_{i\to\infty} d_i < 0$, $(X_n)_{n\geq 0}$ is ergodic.

Lemma B (Kaplan [3]): Under the assumptions of Lemma A, if for some integer $N \geq 0$ and some constants $B \geq 0$, $c \in [0, 1]$ the following two conditions hold, then $(X_n)_{n\geq 0}$ is not ergodic:

i) for all $i \geq N$, $d_i > 0$

ii) for all $i \geq N$, all $\theta \in [c, 1]$, $\theta^i - \Sigma_j P_{ij}\theta^j \geq -B(1-\theta)$.

From (12), it can be seen that $|d_i|$ is finite since

$$|d_i| \leq \lambda + \sum_{n=0}^{\infty} \lambda_n \sum_{j=0}^{i} B_i(j) C_{n+j} \leq 2\lambda + ip.$$

Next, the drift limit is given by the following lemma.

Lemma 1: If C_n has a limit C, finite or not, then $\lim_{i\to\infty} \Sigma_{n=0}^{\infty} \lambda_n \Sigma_{j=0}^{i} B_i(j) C_{n+j} = C$.

Proof of Lemma 1: We consider two separate cases depending on whether C is finite.

1) $C = +\infty$.

Fix $A > 0$ and pick $r \geq 0$ such that $\lambda_r \neq 0$. There exists an integer M such that for all $n \geq M$, $C_n > A$. Fix such an M. Then we have for $i \geq M$

$$\sum_{n=0}^{\infty} \lambda_n \sum_{j=0}^{i} B_i(j) C_{n+j} > \lambda_r \sum_{j=0}^{i} B_i(j) C_{j+r} > \lambda_r A \sum_{j=M}^{i} B_i(j)$$

which terminates the proof, since for any fixed $M \geq 0$

$$\lim_{i\to\infty} \sum_{j=M}^{i} B_i(j) = 1. \tag{13}$$

2) $C < +\infty$.

We have for $i > M$

$$\left| \sum_{n=0}^{\infty} \lambda_n \sum_{j=0}^{i} B_i(j) C_{n+j} - C \right| \leq \sum_{j=0}^{M} B_i(j) \sum_{n=0}^{\infty} \lambda_n |C_{n+j} - C| + \sum_{j=M+1}^{i} B_i(j) \sum_{n=0}^{\infty} \lambda_n |C_{n+j} - C|. \tag{14}$$

Fix $\epsilon > 0$. There exists M such that for all $n > M$, $|C_n - C| < \epsilon/2$. Fix such an M. Then

$$\sum_{j=M+1}^{i} B_i(j) \sum_{n=0}^{\infty} \lambda_n |C_{n+j} - C| < \frac{\epsilon}{2}.$$

Also, if L is an upper bound for C_n

$$\sum_{j=0}^{M} B_i(j) \sum_{n=0}^{\infty} \lambda_n |C_{n+j} - C| \leq 2L \sum_{j=0}^{M} B_i(j) < \frac{\epsilon}{2}$$

for i big enough because (13) holds, which takes care of the first term in (14) and ends the proof of Lemma 1.

Putting together (12) and Lemmas A and 1, we get that 1) if $\lim_{n\to\infty} C_n = +\infty$, then $\lim_{i\to\infty} d_i = -\infty$, and $(X_n)_{n\geq 0}$ is ergodic; and 2) if $\lim_{n\to\infty} C_n = C < +\infty$, then $\lim_{i\to\infty} d_i = \lambda - C$, and $(X_n)_{n\geq 0}$ is ergodic for $\lambda < C$. If $\lambda > C$, we can apply Lemma B and conclude that $(X_n)_{n\geq 0}$ is not ergodic provided that Kaplan's condition ii) holds. This is the purpose of Lemma 2, which is the last step in the proof of Theorem 1.

Lemma 2: If for all $n \geq 1$, $C_n < L$ for some $L \in (0, \infty)$, then Kaplan's condition holds: there exists a constant B, an integer N, and a real $c \in [0, 1]$ such that

$$\theta^i - \sum_j P_{ij}\theta^j \geq -B(1-\theta) \qquad \text{all } i \geq N, \theta \in [c, 1].$$

Proof of Lemma 2: According to [25], it is enough to show that the downward part of the drift, defined as

$$D(i) = -\sum_{k=1}^{i} k P_{i,i-k}$$

is bounded below. From the transition probabilities (8), we get

$$D(i) = -\sum_{k=1}^{i} k \sum_{n=0}^{\infty} \lambda_n \sum_{j=k}^{i} B_i(j) \epsilon_{n+j,n+k}$$

which can also be put in the form

$$D(i) = -\sum_{j=1}^{i} B_i(j) \sum_{n=0}^{\infty} \lambda_n \sum_{k=1}^{j} k \epsilon_{n+j,n+k}$$

from which it follows that

$$D(i) \geq -\sum_{j=1}^{i} B_i(j) \sum_{n=0}^{\infty} \lambda_n C_{n+j} \geq -L.$$

□

Notice that in the proof of Theorem 1 (and this also holds for Theorem 2 below), the exact expression (7) for $B_i(j)$ is never used. The only requirements are that $(B_i(j))_{0\leq j\leq i}$ is a probability distribution, and that (13) holds. Therefore, our results are valid for a larger class of retransmission policies than was first assumed. For example, there could be K priority groups, each with a different retransmission probability.

Although Theorem 1 is quite general, in many practical cases, the reception matrix has a very simple structure and the stability region can be obtained with virtually no computations. This happens for instance in radio networks with capture where all is needed is the limit of the second column of the matrix, or also in the simple case where above a certain collision size N, the transmission is too garbled for the receiver to be able to decode anything correctly, so that $C_n = 0$ for $n > N$.

This last example is a particular case of a noteworthy feature of Theorem 1, namely that the stability region does not depend on any finite number of rows of the reception matrix. In fact, any number of modifications of the matrix that leaves $\lim_{n\to\infty} C_n$ unchanged does not affect the stability region. Although this may be surprising at first sight, it can be intuitively explained by the fundamental instability of the collision channel: unless the

receiver is perfect (all ϵ_{nn} equal to 1), the backlog will eventually exceed any prefixed value with probability one, thus it is the limit of C_n that determines the stability region.

The stability region is also unchanged if the first transmission of packets is delayed. If new packets are backlogged, that is, transmitted for the first time with probability p in each slot following their arrival (this transmission rule appears in the literature as controlled-access or delayed first transmission); the drifts become $d_i = \lambda - \sum_{j=1}^{i} B_i(j)C_j$ for $i \geq 1$, and from Lemmas 1 and 2 the ergodicity region remains the same.

If C_n does not have a limit, Theorem 1 does not give the stable throughput of the system. Even though in almost all practical cases, and indeed in all the examples of Section II, C_n does have a limit, it is conceptually interesting to examine the case when $\liminf_{n\to\infty} C_n \neq \limsup_{n\to\infty} C_n$. It is worth pointing out that adding constraints as strong as the following on the reception matrix still does not imply that C_n has a limit:

i) $(\epsilon_{n0})_{n\geq 1}$ is nondecreasing

ii) $(\epsilon_{nk})_{n\geq k}$ is nonincreasing for all $k \geq 1$

iii) $\epsilon_{nk} \geq \epsilon_{n,k+1}$ for $n \geq 2$, $1 \leq k \leq n-1$

Although the counterexamples we have been able to build are somewhat contrived. Notice that conditions i) and ii) above imply that each column has a limit $\alpha_k = \lim_{n\to\infty} \epsilon_{nk} (k \geq 0)$, which is very likely to happen in practice. In any case, Theorem 2 below still gives some information on the stability region, although the exact result requires in general the complete knowledge of the sequence $(C_n)_{n\geq 1}$. In fact, given any nonnegative numbers $\alpha < \gamma$ β, one can construct a reception matrix with nth row average C_n such that:

i) $\liminf_{n\to\infty} C_n = \alpha$

ii) $\limsup_{n\to\infty} C_n = \beta$

and such that the maximum stable throughput is γ.

Theorem 2: The system is stable for $\lambda < \liminf_{n\to\infty} C_n$ and unstable for $\lambda > \limsup_{n\to\infty} C_n$.

Proof:

a) If $\lambda < \liminf_{n\to\infty} C_n$, then $(X_n)_{n\geq 0}$ is ergodic.

If $\liminf_{n\to\infty} C_n = +\infty$, then $\lim_{n\to\infty} C_n = +\infty$, and the result has already been proved, so assume that $\liminf_{n\to\infty} C_n$ is finite. From Lemma A, it is enough to prove that for all $\epsilon > 0$, there exists N such that

$$d_i < \lambda - \liminf_{n\to\infty} C_n + \epsilon \qquad \text{all } i \geq N.$$

Recall from (12) that we have

$$d_i = \lambda - \sum_{j=0}^{\infty} \lambda_n \sum_{j=0}^{i} B_i(j) C_{n+j}. \tag{15}$$

So it is only needed to prove that for all $\epsilon > 0$ there exists N such that

$$\sum_{n=0}^{\infty} \lambda_n \sum_{j=0}^{i} B_i(j) C_{n+j} > \liminf_{n\to\infty} C_n - \epsilon \qquad \text{all } i \geq N.$$

Now by definition there exists M such that for all $k \geq M$:

$$C_k > \liminf_{n\to\infty} C_n - \epsilon$$

and therefore for all $i > M$:

$$\sum_{n=0}^{\infty} \lambda_n \sum_{j=0}^{i} B_i(j) C_{n+j} > (\liminf_{n\to\infty} C_n - \epsilon) \sum_{j=M}^{i} B_i(j)$$

which completes the proof since (13) holds.

b) If $\lambda > \limsup_{n\to\infty} C_n$, then $(X_n)_{n\geq 0}$ is not ergodic.

Since λ is finite, in this case $\limsup_{n\to\infty} C_n$ is necessarily finite. Therefore, $(C_n)_{n\geq 1}$ is bounded and from Lemma 2, Kaplan's condition holds. Thus, it is enough to show that for all $\epsilon > 0$, there exists N such that

$$d_i > \lambda - \limsup_{n\to\infty} C_n - \epsilon \qquad \text{all } i \geq N.$$

From (15), we only need to show

$$\sum_{n=0}^{\infty} \lambda_n \sum_{j=0}^{i} B_i(j) C_{n+j} < \limsup_{n\to\infty} C_n + \epsilon \qquad \text{all } i \geq N.$$

Since there exists M such that for all $k \geq M$

$$C_k < \limsup_{n\to\infty} C_n + \epsilon$$

then if L is an upper bound for C_n, we have for $i \geq M$

$$\sum_{n=0}^{\infty} \lambda_n \sum_{j=0}^{i} B_i(j) C_{n+j} < L \sum_{j=0}^{M-1} B_i(j) + \limsup_{n\to\infty} C_n + \epsilon$$

from which the result follows, using (13). □

V. Behavior of the Backlog Markov Chain in the Nonergodicity Region

In this section, we further investigate the properties of $(X_n)_{n\geq 0}$ in the case $\lambda > C$, assuming of course that $(C_n)_{n\geq 1}$ has a finite limit. It has been proved in [1] that the backlog Markov chain for the usual slotted Aloha algorithm is transient, but this result cannot be generalized to our model when $\lambda > C$. We give below an example showing that $(X_n)_{n\geq 0}$ can be null recurrent when the mean arrival rate λ belongs to an interval of positive length. The boundary between the null recurrence and the transience regions generally depends in a rather complicated manner on both the reception matrix and the retransmission probability p. We give a sufficient condition for $(X_n)_{n\geq 0}$ to be transient when $\lambda > C$, as well as bounds on the recurrence region.

Consider the reception matrix defined by

$$\epsilon_{nk} = \frac{1}{n^2} \qquad (1 \leq k \leq n)$$

$$\epsilon_{n0} = 1 - \frac{1}{n}$$

for $n \geq 1$. Then $C_n = \sum_{k=1}^{n} k/n^2 = (n+1)/2n$, and $C = 1/2$. Using Lemmas C and D below, we show in [26] that X_n is recurrent for $\lambda < R(p)$ and transient for $\lambda > R(p)$, where $R(p)$ is a function of the retransmission probability p and is given by

$$R(p) = \frac{1}{p} + \frac{(1-p)}{p^2} \ln(1-p) \qquad (0 < p < 1)$$

$$R(1) = 1.$$

It is easily seen that $R(p)$ is an increasing function of p for $p \in]0, 1[$ with extrema $\lim_{p\to 0} R(p) = 1/2$ and $\lim_{p\to 1} R(p) = 1$. Fig. 3 summarizes the behavior of X_n for this example.

It is somehow surprising to see that in this case, as well as in all

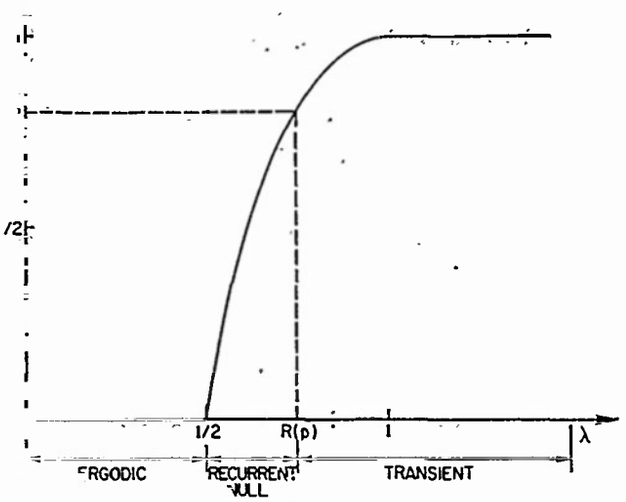

Fig. 3. Transience and ergodicity regions as a function of the retransmission probability when $\epsilon_{nk} = 1/n^2$.

The other examples we have computed, the recurrence region becomes larger as p increases. Intuitively, the recurrence of X_n when $\lambda > C$ seems to be due to the fact that transitions from any state i to 0 (or to some fixed integer k_0) are possible and that the probability of such an event, P_{i0} (or P_{ik_0}), goes to zero slowly with i. It can be checked that these probabilities are increasing functions of p when i is large enough.

Transience is ensured for $\lambda > C$ if the supremum of the elements of the kth-column goes to zero faster than k^2. This condition holds for all the examples in Section II, as well as for many real life cases, due to the practical limitations on the receiver capabilities. In particular, it is always verified if the reception matrix has only a finite number of nonzero columns (or equivalently, if the backlog Markov chain has uniformly bounded downwards transitions, as defined in [3]) which happens, for instance, if there is capture. Note that the proof of Theorem 3 below is of course valid for the conventional collision channel, and in this case becomes somewhat simpler than the proof in [1].

Theorem 3: If $\lim_{k\to\infty} k^2 \sup_{n\geq k} \epsilon_{nk} = 0$, then $(X_n)_{n\geq 0}$ is transient for $\lambda > C$.

Because of the complexity and lack of structure of the one-step transition probabilities (8), few results on the recurrence and transience of Markov chains can be applied to our model. Before proving Theorem 3, let us introduce the following two criteria from [27].

Lemma C: Let $(X_n)_{n\geq 0}$ be an irreducible and aperiodic Markov chain, having as state space the set of nonnegative integers, and with one-step transition probability matrix $P = (P_{ij})$. $(X_n)_{n\geq 0}$ is recurrent if and only if there exists a sequence $(y_n)_{n\geq 0}$ such that

1) $\lim_{i\to\infty} y_i = +\infty$

2) for some integer $N > 0$ $\sum_{j=0}^{\infty} y_j P_{ij} \leq y_i$ all $i \geq N$.

We will only use the sufficiency part, which has also been proved in [24].

Lemma D: With the same assumptions as in Lemma C, $(X_n)_{n\geq 0}$ is transient if and only if there exists a sequence $(y_n)_{n\geq 0}$ such that

1) $(y_n)_{n\geq 0}$ is bounded

2) for some integer $N > 0$ $\sum_{j=0}^{\infty} y_j P_{ij} \leq y_i$ all $i \geq N$

3) for some $k \geq N$ $y_k < y_0, \cdots, y_{N-1}$.

Sufficiency under the additional constraints $y_i > 0$ and $\lim_{i\to\infty} y_i = 0$ has also been proved in [28]. Also, the sufficiency parts of both lemmas are an immediate consequence of [29, Theorems 5 and 6] together with the results in [30].

Proof of Theorem 3: We use Lemma D with $y_n = 1/(n+1)^\theta$, $\theta \in]0, 1[$. We have

$$\sum_j P_{ij} y_j \leq y_i \Leftrightarrow \sum_{k=1}^{i} (y_{i-k} - y_i) P_{i,i-k} + \sum_{k=1}^{\infty} (y_{i+k} - y_i) P_{i,i+k} \leq 0 \quad (16)$$

and

$$(i+1)^{1+\theta} \sum_{k=1}^{i} (y_{i-k} - y_i) P_{i,i-k} + (i+1)^{1+\theta} \sum_{k=1}^{\infty} (y_{i+k} - y_i) P_{i,i+k} = D'(i) + U'(i) \quad (17)$$

where we have defined

$$D'(i) = (i+1)^{1+\theta} \sum_{k=1}^{i} \left[\frac{1}{(i+1-k)^\theta} - \frac{1}{(i+1)^\theta} \right] \cdot \sum_{n=0}^{\infty} \lambda_n \sum_{j=k}^{i} B_i(j) \epsilon_{n+j,n+k}$$

$$U'(i) = (i+1)^{1+\theta} \sum_{k=1}^{\infty} \left[\frac{1}{(i+1+k)^\theta} - \frac{1}{(i+1)^\theta} \right] \cdot \sum_{n=0}^{\infty} \lambda_{k+n} \sum_{j=0}^{i} B_i(j) \epsilon_{n+k+j,n}. \quad (18)$$

The drift of X_n at state i can be computed from the transition probabilities (8)

$$d_i = -\sum_{k=1}^{i} k P_{i,i-k} + \sum_{k=1}^{\infty} k P_{i,i+k} = D(i) + U(i) \quad (19)$$

where we have defined

$$D(i) = -\sum_{k=1}^{i} k \sum_{n=0}^{\infty} \lambda_n \sum_{j=k}^{i} B_i(j) \epsilon_{n+j,n+k}$$

$$U(i) = \sum_{k=1}^{\infty} k \sum_{n=0}^{\infty} \lambda_{n+k} \sum_{j=0}^{i} B_i(j) \epsilon_{j+k+n,n}. \quad (20)$$

The idea of the proof is to show that

$$\lim_{i\to\infty} [D'(i) + U'(i)] = -\theta \lim_{i\to\infty} d_i \quad (21)$$

and since it has been proved in Section III that $\lim_{i\to\infty} d_i = \lambda - C$, we will be able to conclude that $(X_n)_{n\geq 0}$ is transient for $\lambda > C$.

1) $\lim_{i\to\infty} [D'(i) + \theta D(i)] = 0$.

From (18) and (20)

$$D'(i)+\theta D(i)=(i+1)\sum_{k=1}^{i}\left[\left(\frac{i+1}{i+1-k}\right)^{\theta}-1-\frac{\theta k}{i+1}\right]\cdot\sum_{n=0}^{\infty}\lambda_n\sum_{j=k}^{i}B_i(j)\epsilon_{n+j,n+k}$$

which is more conveniently written as

$$D'(i)+\theta D(i)=(i+1)\sum_{n=0}^{\infty}\lambda_n\sum_{j=1}^{i}B_i(j)\sum_{k=1}^{j}\cdot\left[\left(\frac{i+1}{i+1-k}\right)^{\theta}-1-\frac{\theta k}{i+1}\right]\epsilon_{n+j,n+k}.$$

This expression is nonnegative since

$$\left(\frac{i+1}{i+1-k}\right)^{\theta}-1-\frac{\theta k}{i+1}>0 \qquad (1\le k\le i).$$

Define $\gamma_k = \sup_{n\ge k}\epsilon_{nk}$. Then

$$0\le D'(i)+\theta D(i)\le(i+1)\sum_{n=0}^{\infty}\lambda_n\sum_{j=1}^{i}B_i(j)\sum_{k=1}^{j}\cdot\left[\left(\frac{i+1}{i+1-k}\right)^{\theta}-1-\frac{\theta k}{i+1}\right]\gamma_{n+k}$$

$$\le(i+1)\sum_{n=0}^{\infty}\lambda_n\sum_{k=1}^{i}\left[\left(\frac{i+1}{i+1-k}\right)^{\theta}-1-\frac{\theta k}{i+1}\right]\gamma_{n+k}.$$

That is

$$D'(i)+\theta D(i)\le x_1(i)+x_2(i) \tag{22}$$

with, assuming for instance that i is odd

$$x_1(i)=(i+1)\sum_{n=0}^{\infty}\lambda_n\sum_{k=1}^{(i+1)/2}\left[\left(\frac{i+1}{i+1-k}\right)^{\theta}-1-\frac{\theta k}{i+1}\right]\gamma_{n+k}$$

$$x_2(i)=(i+1)\sum_{n=0}^{\infty}\lambda_n\sum_{k=(i+3)/2}^{i}\left[\left(\frac{i+1}{i+1-k}\right)^{\theta}-1-\frac{\theta k}{i+1}\right]\gamma_{n+k}. \tag{23}$$

We show that $x_1(i)$ and $x_2(i)$ go to zero independently. Fix $\epsilon > 0$. Define for $0 < x \le i$ the function

$$p_i(x)=\frac{i+1}{x^2}\left[\left(\frac{i+1}{i+1-x}\right)^{\theta}-1\right]-\frac{\theta}{x}.$$

It is easily proved that for each $i \ge 1$, $p_i(x)$ is a positive nondecreasing function of x. Also

$$p\left(\frac{i+1}{2}\right)=\frac{1}{i+1}[4(2^{\theta}-1)-2\theta]=\frac{A}{i+1}$$

where A is a positive constant depending only on θ. From (23)

$$x_1(i)=\sum_{n=0}^{\infty}\lambda_n\sum_{k=1}^{(i+1)/2}k^2p_i(k)\gamma_{n+k}\le\frac{A}{i+1}\sum_{n=0}^{\infty}\lambda_n\sum_{k=1}^{n+(i+1)/2}k^2\gamma_{n+k}.$$

If $\lim_{k\to\infty}k^2\gamma_k = 0$, then $\lim_{n\to\infty}1/n\Sigma_{k=1}^{n}k^2\gamma_k = 0$. So we can choose i large enough so that for $n \ge (i+1)/2$, $\Sigma_{k=1}^{n}k^2\gamma_k < n\epsilon$. Then

$$x_1(i)\le\epsilon\frac{A}{i+1}\sum_{n=0}^{\infty}\lambda_n\left(n+\frac{i+1}{2}\right)=\epsilon A\left(\frac{\lambda}{i+1}+\frac{1}{2}\right).$$

Now if we choose i big enough so that for $k > (i+3)/2$, we have $\gamma_k < \epsilon/k^2$, then

$$x_2(i)\le\epsilon\sum_{n=0}^{\infty}\lambda_n\sum_{k=(i+3)/2}^{i}(i+1)\cdot\left[\left(\frac{i+1}{i+1-k}\right)^{\theta}-1-\frac{\theta k}{i+1}\right]\frac{1}{(n+k)^2}$$

$$\le\frac{4\epsilon}{i+3}\sum_{k=(i+3)/2}^{i}\left[\left(\frac{i+1}{i+1-k}\right)^{\theta}-1-\frac{\theta k}{i+1}\right].$$

By bounding the sum in the last equation by integrals, it can be seen that it is upper bounded by a linear function of i.

2) $\lim_{i\to\infty}[U'(i)+\theta U(i)] = 0$.

From (18) and (20)

$$U'(i)+\theta U(i)=\sum_{k=1}^{\infty}(i+1)\left[\left(\frac{i+1}{i+1+k}\right)^{\theta}-1+\frac{\theta k}{i+1}\right]\cdot\sum_{n=0}^{\infty}\lambda_{k+n}\sum_{j=0}^{i}B_i(j)\epsilon_{j+k+n,n}.$$

With a change of variable

$$J'(i)+\theta U(i)=(i+1)\sum_{j=0}^{i}B_i(j)\sum_{n=1}^{\infty}\lambda_n\sum_{k=1}^{n}\cdot\left[\left(\frac{i+1}{i+1+k}\right)^{\theta}-1+\frac{\theta k}{i+1}\right]\epsilon_{n+j,n-k}.$$

By using the following inequalities:

$$0\le\frac{1}{(1+x)^{\theta}}-1+\theta x\le\theta(1+\theta)\frac{x^2}{2} \qquad (x\ge 0,\ 0<\theta<1)\}$$

we get

$$0\le U'(i)+\theta U(i)\le\theta\frac{(1+\theta)}{2}(i+1)\sum_{j=0}^{i}B_i(j)\cdot\sum_{n=1}^{N}\lambda_n\sum_{k=1}^{n}\frac{k^2}{(i+1)^2}\epsilon_{n+j,n-k}$$

$$+\theta(i+1)\sum_{j=0}^{i}B_i(j)\sum_{n=N+1}^{\infty}\lambda_n\sum_{k=1}^{n}\frac{k}{i+1}\epsilon_{n+j,n-k}$$

$$\le\frac{1}{i+1}\sum_{n=1}^{N}n^2\lambda_n+\sum_{n=N+1}^{\infty}n\lambda_n.$$

Fix $\epsilon > 0$. Choose N such that $\Sigma_{N+1}^{\infty}n\lambda_n < \epsilon/2$, and then, N being fixed, choose i large enough so that $1/(i+1)\Sigma_{n=1}^{N}n^2\lambda_n < \epsilon/2$. □

It should be clear at this point that unlike the ergodicity region,

he recurrence region depends in general on the elements of the eception matrix (instead of only the row averages) and on the ransmission probability p. For this reason, an exact expression or the recurrence region seems rather difficult to obtain; onetheless, the method (see [26]) that we used to study the xample in Fig. 3 can be generalized to obtain the following upper nd lower bounds on the recurrence region.

heorem 4: $(X_n)_{n\geq 0}$ is recurrent for $\lambda < L$ and transient for $\lambda > U$, with $L = \max\{l_1, \sup_{0<\theta<1} l_\theta, \sup_{0<\theta<1} l'_\theta\}$ and $U = \min\{u_1, \inf_{0<\theta<1} u_\theta, \inf_{0<\theta<1} u'_\theta\}$ where

$$l_1 = \liminf_{i\to\infty} (i+1) \sum_{n=0}^{\infty} \lambda_n \sum_{j=1}^{i} B_i(j) \sum_{k=1}^{n+j} \ln\left(\frac{i+n+1}{i+n-k+1}\right) \epsilon_{n+j,k}$$

$$l_\theta = \frac{1}{\theta} \liminf_{i\to\infty} (i+1)^{1-\theta} \sum_{n=0}^{\infty} \lambda_n \sum_{j=1}^{i} B_i(j) \sum_{k=1}^{n+j} [(i+n+1)^\theta - (i+n-k+1)^\theta] \epsilon_{n+j,k}$$

$$l'_\theta = \frac{1}{\theta} \liminf_{i\to\infty} (i+1)\,[\ln(i+1)]^{1-\theta} \sum_{n=0}^{\infty} \lambda_n \sum_{j=1}^{i} B_i(j) \sum_{k=1}^{n+j} [[\ln(i+n+1)]^\theta - [\ln(i+n-k+1)]^\theta] \epsilon_{n+j,k}$$

nd

$$u_1 = \liminf_{i\to\infty} (i+1)\,[\ln(i+1)]^2 \sum_{n=0}^{\infty} \lambda_n \sum_{j=1}^{i} B_i(j) \sum_{k=1}^{n+j} \left[\frac{1}{\ln(i+n+2-k)} - \frac{1}{\ln(i+n+2)}\right] \epsilon_{n+j,k}$$

$$u_\theta = \frac{1}{\theta} \liminf_{i\to\infty} (i+1)^{1+\theta} \sum_{n=0}^{\infty} \lambda_n \sum_{j=1}^{i} B_i(j) \sum_{k=1}^{n+j} \left[\frac{1}{(i+n+1-k)^\theta} - \frac{1}{(i+n+1)^\theta}\right] \epsilon_{n+j,k}$$

$$u'_\theta = \frac{1}{\theta} (i+1)\,[\ln(i+1)]^{1+\theta} \sum_{n=0}^{\infty} \lambda_n \sum_{j=1}^{i} B_i(j) \sum_{k=1}^{n+j} \left[\frac{1}{[\ln(i+n+2-k)]^\theta} - \frac{1}{[\ln(i+n+2)]^\theta}\right] \epsilon_{n+j,k}.$$

Ve are assuming that the limits above exist, which indeed appens in most practical cases. The theorem is valid if any of hese limits is infinite. In particular, if $L = +\infty$, then X_n is lways recurrent. Note that usually, it is not necessary to carry out ll the computations, because one of the three terms in the efinition of L is equal to one of the terms in the definition of U, n fact, in most cases, we have $\sup_{0<\theta<1} l_\theta = \inf_{0<\theta<1} u_\theta$ if $0 < p < 1$. and $u_1 = l_1$ if $p = 1$. The proof of Theorem 4 can be found n [26].

References

[1] W. A. Rosenkrantz and D. Towsley, "On the instability of the slotted Aloha multiaccess algorithm," *IEEE Trans. Automat. Contr.*, vol. AC-28, pp. 994–996, 1983.

[2] G. Fayolle, E. Gelenbe, and J. Labetoulle, "Stability and optimal control of the packet switching broadcast channel," *J. Ass. Comput. Mach.*, vol. 24, no. 3, pp. 375–386, 1977.

[3] M. Kaplan, "A sufficient condition for nonergodicity of a Markov chain," *IEEE Trans. Inform. Theory*, vol. IT-25, pp. 470–471, 1979.

[4] D. Bertsekas and R. Gallager, *Data Networks*. Englewood Cliffs, NJ: Prentice-Hall, 1987.

[5] B. Hajek and T. Van Loon, "Decentralized dynamic control of a multiaccess broadcast channel," *IEEE Trans. Automat. Contr.*, vol. AC-27, pp. 559–569, 1982.

[6] B. Hajek, "Hitting-time and occupation-time bounds implied by drift analysis with applications," *Adv. Appl. Prob.*, vol. 14, pp. 502–525, 982.

[7] J. N. Tsitsiklis, "Analysis of a multiaccess control scheme," *IEEE Trans. Automat. Contr.*, vol. AC-32, pp. 1017–1020, 1987.

[8] S. Verdú, "Minimum probability of error for asynchronous Gaussian multiple-access channels," *IEEE Trans. Inform. Theory*, vol. IT-32, pp. 85–96, 1986.

[9] D. Raychaudhuri, "Performance analysis of random access packet-switched code division multiple access systems," *IEEE Trans. Commun.*, vol. COM-29, pp. 895–901, 1981.

[10] A. Polydoros and J. Silvester, "Slotted random access spread-spectrum networks: An analytical framework," *IEEE J. Select. Areas Commun.*, vol. SAC-5, no. 6, pp. 989–1002, 1987.

[11] M. K. Simon, J. K. Omura, R. A. Scholtz, and B. K. Levitt, *Spread-Spectrum Communications*. New York: Computer Science Press, 1985.

[12] N. Abramson, "The throughput of packet broadcasting channels," *EEE Trans. Commun.*, vol. COM-25, pp. 117–128, 1977.

[13] D. J. Goodman and A. A. M. Saleh, "The near/far effect in local Aloha radio communications," *IEEE Trans. Vehicular Technol.*, vol. VT-36, Feb. 1987.

[14] L. G. Roberts, "Aloha packet system with and without slots and capture," *Comput. Commun. Rev.*, no. 5, pp. 28–42, 1975.

[15] C. Namislo, "Analysis of mobile radio slotted Aloha networks," *IEEE . Select. Areas Commun.*, vol. SAC-2, pp. 583–588, 1984.

[16] D. H. Davis and S. A. Gronemeyer, "Performance of slotted Aloha random access with delay capture and randomized time of arrival," *IEEE Trans. Commun.*, vol. COM-28, pp. 703–710, 1980.

[17] J. C. Arnbak and W. van Blitterswijk, "Capacity of slotted Aloha in Rayleigh fading channels," *IEEE J. Select. Areas Commun.*, vol. SAC-5, no. 2, pp. 261–269, 1987.

[18] C. C. Lee, "Random signal levels for channel access in packet radio networks," *IEEE J. Select. Areas Commun.*, vol. SAC-5, no. 6, pp. 1026–1034, 1987.

[19] B. S. Tsybakov, V. A. Mikhailov, and N. B. Likhanov, "Bounds for packet transmission rate in a random multiple access system," *Problemy Peredachi Informatsii*, vol. 19, no. 1, pp. 61–81, 1983.

[20] N. Mehravari, "Collision resolution in random access systems with multiple reception," preprint, 1987.

[21] A. Shwartz and M. Sidi, "Erasure, capture and noise errors in controlled multiple access networks," in *Proc. 25th Conf. Decision Contr.*, Dec. 1986, pp. 1333–1334.

[22] R. E. Kahn, S. A. Gronemeyer, J. Burchfiel, and R. C. Kunzelman, "Advances in packet radio technology," *Proc. IEEE*, vol. 66, pp. 1468–1496, 1978.

[23] W. Szpankowski, "Analysis and stability considerations in a reservation multiaccess system," *IEEE Trans. Commun.*, vol. COM-31, no. 5, pp. 684–692, 1983.

[24] A. G. Pakes, "Some conditions for ergodicity and recurrence of Markov chains," *Oper. Res.*, vol 17, pp. 1058–1061, 1969.

[25] L. I. Sennott, P. A. Humblet, and R. L. Tweedie, "Mean drifts and the non-ergodicity of Markov chains," *Oper. Res.*, vol. 31, pp. 783–789, 1983.

[26] S. Ghez, S. Verdú, and S. C. Schwartz, "Random access communication with multipacket reception capability. I: Stability of Aloha," Dep. Elect. Eng., Princeton Univ., Tech. Rep. ISS. 8748, 1987.

[27] J. F. Mertens, E. Samuel-Cahn, and S. Zamir, "Necessary and sufficient conditions for recurrence and transience of Markov chains, in terms of inequalities," *J. Appl. Prob.*, vol. 15, pp. 848–851, 1978.

[28] C. M. Harris and P. G. Marlin, "A note on feedback queues with bulk service," *J. Assoc. Mach.*, vol. 19, pp. 727–733, 1972.

[29] F. G. Foster, "On the stochastic matrices associated with certain queuing processes," *Ann. Math. Statist.*, no. 24, pp. 355–360, 1953.

[30] P. Vit, "On the equivalence of certain Markov chains," *J. Appl. Prob.*, vol. 13, pp. 357–360, 1976.

[31] S. Ghez, S. Verdú, and S. C. Schwartz, "On decentralized control algorithms for multipacket Aloha," in *Proc. 25th Allerton Conf. on Commun., Contr., and Comput.*, Oct. 1987.

Sylvie Ghez (S'88) was born in [redacted]. She received the engineering diploma from the Ecole Nationale Supérieure des Télécommunications, Paris, in 1984, and the M.S. degree in electrical engineering from Princeton University, Princeton, NJ, in 1987, where she is currently a doctoral candidate.

She is the recipient of the Wallace Memorial Fellowship for the year 1987–1988. Her research interests include multiple access protocols, network routing problems, queueing theory, and applied probability theory.

Sergio Verdú (S'80–M'84) received the Ph.D. degree in electrical engineering from the University of Illinois at Urbana-Champaign in 1984.

Upon completion of his doctorate he joined the faculty of Princeton University, Princeton, NJ, where he is an Assistant Professor of Electrical Engineering. His recent research contributions are in the areas of multiuser communication and information theory, and statistical signal processing.

Dr. Verdú is a recipient of the National University Prize of Spain, the Rheinstein Outstanding Junior Faculty Award of the School of Engineering and Applied Science, Princeton University, and the NSF Presidential Young Investigator Award. He is currently serving as Associate Editor of the IEEE TRANSACTIONS ON AUTOMATIC CONTROL.

Stuart C. Schwartz (S'64–M'66–SM'83) was born in [redacted]. He received the B.S. and M.S. degrees in aeronautical engineering from the Massachusetts Institute of Technology, Cambridge, in 1961, and the Ph.D. degree from the Information and Control Engineering Program, University of Michigan, Ann Arbor, in 1966.

While at M.I.T. he was associated with the Naval Supersonic Laboratory and the Instrumentation Laboratory. From 1961 to 1962 he was at the Jet Propulsion Laboratory, Pasadena, CA, working on problems in orbit estimation and telemetry. He is presently a Professor of Electrical Engineering and Chairman of the department at Princeton University, Princeton, NJ. He served as Associate Dean of the School of Engineering from July 1977 to June 1980. During the academic year 1972–1973, he was a John S. Guggenheim Fellow and a Visiting Associate Professor at the Department of Electrical Engineering, Technion—Israel Institute of Technology, Haifa, Israel. During the academic year 1980–1981, he was a member of the Technical Staff at the Radio Research Laboratory, Bell Telephone Laboratories, Crawford Hill, NJ. His principal research interests are in the application of probability and stochastic processes to problems in statistical communication and system theory.

Dr. Schwartz is a member of Sigma Gamma Tau, Eta Kappa Nu, and Sigma Xi. He has served as an Editor for the *SIAM Journal on Applied Mathematics* and as Program Chairman for the 1986 ISIT.

T-COM/37/1//24901

Performance Analysis of an Asymptotically Quantum-Limited Optical DPSK Receiver

David Brady
Sergio Verdú

Reprinted from
IEEE TRANSACTIONS ON COMMUNICATIONS
Vol. 37, No. 1, January 1989

Performance Analysis of an Asymptotically Quantum-Limited Optical DPSK Receiver

DAVID BRADY, STUDENT MEMBER, IEEE, AND SERGIO VERDÚ, SENIOR MEMBER, IEEE

Abstract—In this paper, we analyze an optical, direct-detection DPSK receiver whose error probability is quantum-limited as the transmitting laser linewidth vanishes. The receiver design is based on a binary equiprobable hypothesis test with doubly stochastic point process observations, the conditional random rates of which depend on the transmitting laser phase noise, which is modeled as a Brownian motion. The receiver structure consists of a simple, delay-and-sum optical preprocessor followed by a photoelectric converter and an integrate-and-dump circuit. Upper and lower bounds on the receiver bit error rate are derived by developing bounds on the conditional rates of the point process, and it is shown that the error probability bounds converge to the true value as the transmitting laser linewidth decreases. Bounds on the power penalty are computed for parameters corresponding to existing semiconductor injection lasers, and are seen to be less than the limiting power penalty for the balanced DPSK receiver.

I. Introduction

In differential phase shift keying (DPSK), information is conveyed by the carrier phase in the current symbol interval relative to that in the previous interval. While less efficient than phase shift keying (PSK), DPSK is less sensitive to large phase noise amplitudes by utilizing phase noise correlation in adjacent symbol intervals. Demodulation in conventional (radio frequency) DPSK systems can be achieved by multiplying the total received scalar field by a delayed version of itself, followed by integration over a symbol period [1]. However, due to the lack of efficient optical multipliers and sharp filters, this receiver structure is not yet feasible at optical frequencies. An alternative solution is to heterodyne the received optical signal to the microwave frequency range, and employ a conventional demodulation scheme. This heterodyne DPSK receiver has been analyzed previously [2]. While incurring a 3 dB loss inherent to the heterodyning operation, this receiver avoids the need to count photons, which may introduce an appreciable loss in some existing avalanche photodiodes [3].

In this paper, we assume an ideal, photon-counting device and concentrate on the design and analysis of a direct-detection DPSK receiver. Performance is measured by the *power penalty*, which is the ratio of the transmitted optical power required to achieve a given bit error rate to that required by a receiver whose power requirement is determined solely by the statistical nature of the photodetection process. Thus, the power penalty is a measure of demodulation efficiency, and a receiver with 0 dB power penalty is described as *quantum limited*. In [2], a balanced, direct-detection DPSK receiver was found to have a power penalty of at least 3 dB. We analyze in this paper a DPSK receiver whose power penalty is 0 dB for a transmitting laser with no phase noise, and less than 3 dB at 10^{-9} bit error rate (BER) for existing semiconductor injection lasers.

The remainder of this paper is organized as follows: In Section II, we formulate a DPSK receiver from a binary hypothesis test with point process observations. The randomness of the rates of the point process under each hypothesis is due to the transmitting laser phase noise, which is modeled as a Brownian motion. It is shown why the optimum receiver is not feasible, and then attention is restricted to simple strategies that use photon counting only. The proposed DPSK receiver implements a suboptimum binary hypothesis test preceded by a delay-and-sum optical preprocessor. In Section III, we analyze the receiver performance as a function of system parameters. The exact error probability expression depends on the moment generating function (mgf) of a functional of the laser phase noise sample path. Since this mgf appears to be intractable, the receiver error probability is bounded by using an alternative functional whose mgf is computable.

II. Optical DPSK Receiver

In binary DPSK modulation, the transmitted scalar field is amplitude-modulated by a bit stream derived from the information sequence. Denoting the information sequence as $\{\cdots b_{-1}, b_0, b_1 \cdots\}$ where $b_n \in \{-1, 1\}$ is the information bit in the time interval $nT \le t < (n+1)T$, we compose a sequence $\{\cdots a_{-1}, a_0, a_1 \cdots\}$, $a_n \in \{-1, 1\}$ from the relation $a_{n-1}a_n \triangleq b_n$ to amplitude-modulate the transmitted scalar electric field. Under the assumptions of spatial homogeneity and distortionless transmission, the transmitted (and received) scalar electric fields may be described as

$$s(t) = a_n A \cos(\nu t + \theta_t) = \sqrt{2}\,\mathrm{Re}\left[\frac{a_n A}{\sqrt{2}} e^{j\theta_t} e^{j\nu t}\right], \quad nT \le t < (n+1)T \quad (1)$$

where the Brownian motion $\{\theta_t,\ t \in \mathbb{R}\}$ models the transmitted laser phase noise and ν is the carrier frequency. Note that in the absence of laser phase noise the transmitted optical energy is $\zeta \triangleq A^2T/2$ photons per bit.

The decoding from $\{a_l\}$ to $\{b_l\}$ is performed in the same operation that compares the received signal in $[nT, (n+1)T)$ to the reference signal. This suggests the demodulation scheme shown in Fig. 1. The optical signal described in (1) is divided by a half-silvered mirror into two signals of equal power, one of which is delayed by the symbol period T, and then added to the other. The resultant optical signal is incident on the photodetector in Fig. 1 and is given by

$$r(t) = \frac{1}{\sqrt{2}}\{a_n A \cos(\nu t + \theta_t) + a_{n-1} A \cdot \cos([\nu t + \theta_t] - [\nu T + \Delta\theta_t])\} = \sqrt{2}\,\mathrm{Re}\,[E_t e^{j\nu t}] \quad (2)$$

Paper approved by the Editor for Synchronization and Optical Detection of the IEEE Communications Society. Manuscript received September 17, 1987; revised April 25, 1988. This work was supported by the U.S. Office of Naval Research under Grant N00014-87-K-0054. This paper was presented at the 1988 International Symposium on Information Theory, Kobe, Japan.

The authors are with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544.

IEEE Log Number 8824901.

0090-6778/89/0100-0046$01.00

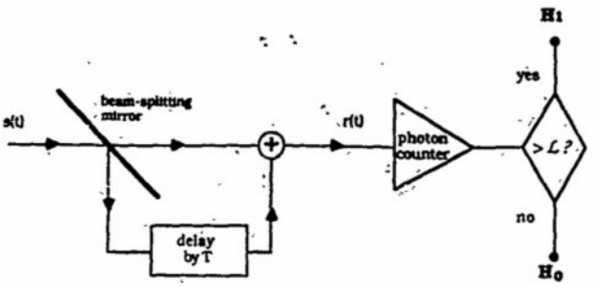

Fig. 1. Proposed optical DPSK receiver.

where $E_t \triangleq A/2e^{j\theta_t}\{a_n + a_{n-1}e^{-j(\nu T+\Delta\theta_t)}\}$ is the complex envelope of the electric field incident on the photodetector, and $\Delta\theta_t$ is an increment of Brownian motion given by

$$\Delta\theta_t \triangleq \theta_t - \theta_{t-T}. \tag{3}$$

The remaining receiver structure will be based on the electron point process observations at the output of the photodetector. This process may be modeled as a doubly stochastic point process with count N_t, occurrence times $\{W_1, W_2, \cdots, W_{N_t}\}$, and rate $\lambda_t = \alpha|E_t|^2 + d$ where α and d are parameters of the photodetector [4]. It should be pointed out that it is equivalent to express the rate as $\lambda_t = \alpha r^2(t) + d$, and to ignore the double frequency terms. Although the following results may easily be extended for arbitrary values of α and d, we assume in the following that the photoelectric conversion is ideal, i.e., $\alpha = 1$ and $d = 0$. Under these assumptions the rate of the electron departure process for $0 \le t < T$ is

$$\lambda_t = \frac{\zeta}{T}[1 + b_0 \cos \Delta\theta_t], \quad 0 \le t < T \tag{4}$$

where we have assumed $\nu T = 0 \bmod 2\pi$. The hypothesis pair for the interval $[0, T)$ may be described as

N is a doubly stochastic counting process with intensity:

$$H_0 : \lambda_t = \lambda_t^{(0)} \triangleq \frac{\zeta}{T}(1 - \cos \Delta\theta_t) \quad 0 \le t < T$$

$$H_1 : \lambda_t = \lambda_t^{(1)} \triangleq \frac{\zeta}{T}(1 + \cos \Delta\theta_t) \quad 0 \le t < T \tag{5}$$

where ζ is the transmitted optical energy in photons per bit as defined earlier. If we define $\hat{\lambda}_t^{(i)} \triangleq E[\lambda_t^{(i)} | \{N_\sigma; 0 \le \sigma < t\}]$, then the test which minimizes the probability of error is [5]

$$\sum_{t=W_i}^{N_T} \log \frac{\hat{\lambda}_t^{(1)}}{\hat{\lambda}_t^{(0)}}\bigg|_{t=W_i} - \int_0^T \hat{\lambda}_t^{(1)} - \hat{\lambda}_t^{(0)}\, dt \underset{H_0}{\overset{H_1}{\gtrless}} 0. \tag{6}$$

The formulation of this log-likelihood ratio has an interesting, two part structure. First derive the minimum mean squared error causal estimates of the intensities given the observation under each hypothesis, and then solve a binary hypothesis testing problem with observations from a nonhomogeneous Poisson process whose conditional rates are these estimates. This fact has commonly been referred to as the *separation theorem of detection*, and motivates the use of suboptimal estimates in hypothesis testing with doubly stochastic point process observations [4], [5].

Unfortunately, the explicit structure of (6) is unknown due to the difficulty in evaluating the conditional estimates $\hat{\lambda}_t^{(i)}$. One approach to this problem is to replace the optimum estimates $\hat{\lambda}_t^{(i)}$ with suboptimum estimates. In this paper, we propose the suboptimum estimates $E[\lambda_t^{(i)}]$. A justification of this approach is the following. Denoting B_l as the laser linewidth in Hertz, $\gamma \triangleq 1/B_l T$, and choosing x_0 such that $1 - e^{-\pi/\gamma} < x_0 < e^{-\pi/\gamma}$, we have

$$P[\exists t \in [0, T] : |\cos \Delta\theta_t - E[\cos \Delta\theta_t]| > x_0] \le 2e^{-(\gamma/32\pi)[\cos^{-1}(e^{-(\pi/\gamma)} - x_0)]^2}. \tag{7}$$

This upper bound follows from Lemma 1 in Section III. For $\gamma = 1000$, and $x_0 = 20$ percent of $E[\cos \Delta\theta_t]$, the left-hand side of (7) is less than 0.03. For $\gamma = 1500$ and $x_0 = 20$ percent of $E[\cos \Delta\theta_t]$, the probability is less than 0.003. So the probability that a sample path of the intensity deviates from the mean in a symbol interval by more than 20 percent is bounded above by a small number for reasonable values of γ. Based on this argument, we employ the (constant) mean $E[\cos \Delta\theta_t]$ to estimate the process $\{\cos \Delta\theta_t, t \in [0, T]\}$, and the suboptimal estimates of the conditional rates follow directly. The advantage to assuming a constant estimate is that the test assumes homogeneous Poisson point process observations, and the decision strategy is very simple: compare the photon count to a threshold. By (7) we can expect this strategy to be optimum for large γ although it may be far from optimum for low γ.

A suboptimal receiver design is specified below. Since $\{\theta_t, t \in \mathbb{R}\}$ is a Brownian motion with zero drift and diffusion coefficient $\sqrt{2\pi B_l}$, $\{\Delta\theta_t, t \in \mathbb{R}\}$ is a stationary Gaussian random process with zero mean and autocorrelation function

$$R(\tau) = \begin{cases} \dfrac{2\pi}{\gamma}\left[1 - \dfrac{|\tau|}{T}\right], & |\tau| < T \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

The means of $\lambda_t^{(i)}$ are easily computed

$$E[\lambda_t^{(0)}] = \frac{\zeta}{T}(1 - e^{-\pi/\gamma})$$

$$E[\lambda_t^{(1)}] = \frac{\zeta}{T}(1 + e^{-\pi/\gamma}). \tag{9}$$

Employing these suboptimal estimates in (6), a suboptimal hypothesis test is

$$N_T \underset{H_0}{\overset{H_1}{\gtrless}} \left\lfloor \frac{2\zeta e^{-\pi/\gamma}}{\log\left[\dfrac{1 + e^{-\pi/\gamma}}{1 - e^{-\pi/\gamma}}\right]} \right\rfloor \tag{10}$$

where $\lfloor \cdot \rfloor$ indicates the greatest integer function. Due to the stationarity of the laser phase noise increment, the suboptimal test is not a function of the photon arrival times, and the remaining portion of the DPSK receiver need only count the number of received photons in the interval $[0, T]$. The entire suboptimal receiver structure is shown in Fig. 1. Note that the threshold obtained in (10) is not the threshold that minimizes the error probability of a test which uses the statistic N_T. The optimum threshold is also a function of the optical power and γ, and is closely approximated as a by-product of the error probability analysis of Section III.

It is worthwhile to compare this DPSK receiver to the balanced receiver analyzed in [2]. The balanced receiver divides the received optical signal into four streams of equal intensity, two of which are delayed by the symbol period T. A delayed stream is added to a nondelayed stream in the same way as in the proposed receiver, except at one-half the power. The remaining pair of optical streams are subtracted. The outputs of the summer and subtractor are each input to photoelectric converters, and the resulting counts are subtracted. The resulting hypothesis test requires no threshold setting. That is, the difference of the electron counts is compared to zero, and the receiver is independent of γ.

However, in using this information about the laser phase noise, the proposed DPSK receiver requires one-half the transmitted optical power of the balanced DPSK receiver to achieve the same error rate. This claim will be verified in Section III. This comparison assumes that the three-port beam combiners used in both receivers are lossless. It should be noted that if a Mach–Zehnder interferometer is used as the beam combiner [6], then the proposed DPSK detector uses only one of the two output port beams, and the energy of the signal incident on the photodetector is one-half of that assumed in the analysis of this paper. In this case, the performance of our receiver is asymptotically equivalent to that of the balanced detector in [2], implemented using both output ports of a Mach–Zehnder interferometer [7].

As suggested above, the parameter γ is central in the analysis of optical communication systems employing coherent light with nonzero linewidth. It characterizes the performance degradation due to the transmitting laser phase jitter relative to the symbol rate. For fixed laser linewidth, the effect of the phase noise on system performance is less pronounced as the symbol rate increases, as reflected by an increase of γ. Typically, $\gamma \in [50, 1600]$, which follows from $B_l \in$ [6 MHz, 20 MHz] for semiconductor injection lasers [8], and symbol rates from 1 to 10 Gbits/s.

III. Performance Analysis

In this section, we characterize the performance of the proposed DPSK receiver. We show that as $\gamma \to \infty$ the probability of error is the quantum limit of optical communications. We then derive upper and lower bounds on the probability of error for arbitrary γ and show that these bounds converge to the true value as $\gamma \to \infty$. Finally, we present Monte Carlo simulation results of the receiver performance and compare them to the error probability bounds.

We begin by showing that the performance of the proposed DPSK receiver is quantum limited as $\gamma \to \infty$. In this case, the transmitting laser is ideal and it is easy to see that (10) becomes

$$N_T \underset{H_0}{\overset{H_1}{\gtrless}} 0. \tag{11}$$

As $\gamma \to \infty$, the rates under each hypothesis are deterministic and N_T is an unconditionally Poisson random variable under each hypothesis. If we define $\Lambda_l \triangleq \int_0^T \lambda_t^{(l)}\, dt$, $l \in \{0, 1\}$ then the probability of error is

$$P[\text{error}] = \frac{1}{2} P[N_T = 0 | H_1] = \frac{1}{2} e^{-\Lambda_1}. \tag{12}$$

Since $\Lambda_1 = 2\zeta$ for an ideal transmitting laser, (12) becomes

$$P[\text{error}] = \frac{1}{2} e^{-2\zeta} \tag{13}$$

which is known as the quantum limit. Thus, the receiver performance is quantum limited as the transmitting laser linewidth goes to zero.

Next, we consider the error probability for finite values of γ. It is convenient to define the moment generating function $M_l(v) \triangleq E[e^{v\Lambda_l}]$ and to let $N(\Lambda)$ denote a Poisson random variable with mean Λ. Conditioned on the rate and the hypothesis, the observation process is a Poisson point process. Therefore, the probability of error under H_1 may be found by first conditioning on $\{\Delta\theta_t, 0 \le t \le T\}$

$$\begin{aligned} P[\text{error}|H_1] &= E[P[N(\Lambda_1) \le l | \{\Delta\theta_t, 0 \le t \le T\}]] \\ &= E\left[\sum_{k=0}^{l} \frac{(\Lambda_1)^k}{k!} e^{-\Lambda_1}\right] \\ &= \sum_{k=0}^{l} \frac{1}{k!} \frac{d^k}{dv^k} M_1(v)\big|_{v=-1}. \end{aligned} \tag{14}$$

where l is an arbitrary nonnegative integer threshold, and the last line follows from an application of the bounded convergence theorem. By a similar argument we have under H_0

$$P[\text{error}|H_0] = 1 - \sum_{k=0}^{l} \frac{1}{k!} \frac{d^k}{dv^k} M_0(v)\big|_{v=-1}. \tag{15}$$

It appears that there is no closed-form expression for the foregoing moment generating functions when $\{\Delta\theta_t, 0 \le t \le T\}$ is a T increment of Brownian motion. In the remainder of this section, we consider upper and lower bounds to (14) and (15). Based on the fact that the mgf of $1/2T \int_0^T (\Delta\theta_t)^2\, dt$ is computable (see the Appendix), our approach is to find upper and lower bounds on the error probability based on quadratic bounds of $\cos x$

$$1 - \frac{x^2}{2} \le \cos x \le f_a(x) \triangleq \begin{cases} 1 - a\dfrac{x^2}{2} & \text{if } |x| < x_a \\ 1 & \text{otherwise} \end{cases} \tag{16}$$

where $0 < a < 1$ and x_a is the smallest positive real number such that $1 - ax_a^2/2 = \cos x_a$. If each cosine in the expressions for Λ_l is replaced by the upper bounding function, the corresponding rates are further apart, hence it is easier to discriminate between them and a lower bound on the error probability is obtained. Analogously, an upper bound is obtained if each cosine is replaced by the lower bounding function. More precisely, the cumulative distribution function (cdf) of a Poisson random variable is a strictly decreasing function of its mean, that is, $P[N(\Lambda) \le l]$ is decreasing in Λ. If the error probability under each hypothesis is found by first conditioning on $\{\Delta\theta_t, 0 \le t \le T\}$, then the conditional probabilities are Poisson cdf's, and substituting the bounds from (16) in the expressions for Λ_l will yield upper and lower bounds for the conditional probabilities as well as the unconditioned probabilities. Denoting Λ_l^U as the bound on Λ_l which yields an upper bound to the error probability on H_l, and Λ_l^L as the bound on Λ_l which yields a lower bound to the error probability under H_l, we have

$$\begin{aligned} \Lambda_0^U &= \frac{\zeta}{2T} \int_0^T (\Delta\theta_t)^2\, dt \\ \Lambda_1^U &= 2\zeta - \Lambda_0^U \\ \Lambda_0^L &= \frac{\zeta}{T} \int_0^T 1 - f_a(\Delta\theta_t)\, dt \\ \Lambda_1^L &= 2\zeta - \Lambda_0^L. \end{aligned} \tag{17}$$

As seen by (14), (15), and (17) computation of these bounds requires the moment generating functions of Λ_0^L and Λ_0^U. The following two lemmas loosen two of the bounds so that we require only $\Phi(v) \triangleq E[e^{v\Lambda_0^U}]$, which is derived in the Appendix. Lemma 1 quantifies the fact that the probability that a T length sample path of $\Delta\theta$ deviates far from the origin decreases with increasing γ. Lemma 2 uses this fact in deriving a lower bound on the error probability under H_0, which does not depend on the mgf of Λ_0^L.

Lemma 1:[1] Let $x \in \mathbb{R}_+$. Then

$$P[\exists t \in [0, T] : |\Delta\theta_t| > x] \le 4Q\left[\frac{x}{4}\sqrt{\frac{\gamma}{4}}\right].$$

[1] $Q[x] \triangleq 1/\sqrt{2\pi} \int_x^\infty e^{-t^2/2}\, dt$.

Proof: Let $\{W_t, t \in \mathbb{R}_+\}$ be a Wiener process, and $\kappa \triangleq x/2\sqrt{2\pi B_l}$. For $v \in \mathbb{R}_+$, we define the stopping times $T_v \triangleq \inf\{t: W_t > v\}$ and $T_{-v} \triangleq \inf\{t: W_t < -v\}$. Then,

$$
\begin{aligned}
&P[\exists t \in [0, T] : |\Delta\theta_t| > x] \\
&\quad = P[\exists t \in [0, T] : |\theta_t - \theta_{t-T}| \geq x] \\
&\quad \leq P\left[\exists t \in [0, 2T] : |\theta_{-T+t} - \theta_{-T}| > \frac{x}{2}\right] \\
&\quad = P\left[\exists t \in [0, 2T] : |\theta_t - \theta_o| > \frac{x}{2}\right] \qquad (18) \\
&\quad = P\left[\exists t \in [0, 2T] : |W_t| > \frac{x}{2\sqrt{2\pi B_l}}\right] \\
&\quad = P[\min(T_\kappa, T_{-\kappa}) < 2T] \\
&\quad \leq P[T_\kappa < 2T] + P[T_{-\kappa} < 2T] \qquad (19) \\
&\quad = 2P[T_\kappa < 2T] \qquad (20) \\
&\quad = 2\{P[T_\kappa < 2T, W_{2T} \geq \kappa] + P[T_\kappa < 2T, W_{2T} < \kappa]\} \\
&\quad = 4P[T_\kappa < 2T, W_{2T} \geq \kappa] \qquad (21) \\
&\quad = 4P[W_{2T} \geq \kappa] \\
&\quad = 4Q\left[\frac{x}{4}\sqrt{\frac{\gamma}{\pi}}\right].
\end{aligned}
$$

Equation (18) follows from the Markov property of the Brownian motion, (19) from the union bound, (20) from the fact that $-W$ has the same probability law as W, and (21) from the reflection principle of the Wiener process. ■

It should be pointed out that a tighter upper bound is possible by using the first passage times for the process $\{\Delta\theta_t, t \in [0, T]\}$, whose distributions are known [9], or by strengthening the inequality in (19) using the first passage time of Brownian motion out of a symmetric interval about the origin [10], [11]. However, the easy upper bound used in the proof suffices for our needs.

Lemma 2: Let l be an arbitrary, nonnegative integer. Lower bounds on the conditional error probabilities are

$$
\begin{aligned}
&1 - 4Q\left[\frac{x_a}{4}\sqrt{\frac{\gamma}{\pi}}\right] - E[P[N(a\Lambda_0^U) \leq l | \{\Delta\theta_t, 0 \leq t \leq T\}]] \\
&\quad \leq E[P[N(\Lambda_0) > l | \{\Delta\theta_t, 0 \leq t \leq T\}]]
\end{aligned}
$$

where $1 - ax^2/2 \geq \cos x \ \forall |x| \in [0, x_a]$, and

$$
\sum_{k=0}^{l} \frac{1}{k!}(2\zeta)^k e^{-2\zeta} \leq E[P[N(\Lambda_1) \leq l | \{\Delta\theta_t, 0 \leq t \leq T\}]].
$$

Proof: Let 1_A be the indicator function of event A. Then

$$
\begin{aligned}
&E[P[N(\Lambda_0) > l | \{\Delta\theta_t, 0 \leq t \leq T\}]] \\
&\quad = 1 - E[P[N(\Lambda_0) \leq l | \{\Delta\theta_t, 0 \leq t \leq T\}]] \\
&\quad = 1 - E[P[N(\Lambda_0) \leq l | \{\Delta\theta_t, 0 \leq t \leq T\}] \\
&\qquad \cdot \{1_{\{|\Delta\theta_t| \leq x_a, \forall t \in [0,T]\}} + 1_{\{\exists t \in [0,T] : |\Delta\theta_t| > x_a\}}\}] \\
&\quad \geq 1 - E[P[N(a\Lambda_0^U) \leq l | \{\Delta\theta_t, 0 \leq t \leq T\}] \\
&\qquad \cdot 1_{\{|\Delta\theta_t| \leq x_a, \forall t \in [0,T]\}}] \\
&\qquad - P[\exists t \in [0, T] : |\Delta\theta_t| > x_a] \\
&\quad \geq 1 - E[P[N(a\Lambda_0^U) \leq l | \{\Delta\theta_t, 0 \leq t \leq T\}]] \\
&\qquad - P[\exists t \in [0, T] : |\Delta\theta_t| > x_a]
\end{aligned}
$$

where $1 - ax^2/2 - \cos x \geq 0 \ \forall |x| \in [0, x_a]$.

Applying Lemma 1, we have proved the first bound of this claim. To prove the second bound, we recognize the fact that $\Lambda_1 < 2\zeta$, and that the cdf of $N(\Lambda)$ is a monotonically decreasing function of Λ. ■

As a direct result of Lemma 2 a lower bound on the total probability of error is

$$
\begin{aligned}
&1 - 4Q\left[\frac{x_a}{4}\sqrt{\frac{\gamma}{\pi}}\right] + \sum_{k=0}^{l} \frac{1}{k!}(2\zeta)^k e^{-2\zeta} \\
&\qquad - \sum_{k=0}^{l} \frac{a^k}{k!} \frac{d^k}{dv^k} \Phi(v)\Big|_{v=-a} \leq 2P[\text{error}] \qquad (22)
\end{aligned}
$$

where $\Phi(v)$ is as defined earlier, and l is an arbitrary nonnegative integer. From (17) as well the fact that the Poisson cdf is a decreasing function of the mean, an upper bound on the total error probability is

$$
2P[\text{error}] \leq 1 + \sum_{k=0}^{l} \frac{1}{k!} \frac{d^k}{dv^k} [e^{2\zeta v}\Phi(-v) - \Phi(v)]_{v=-1}. \qquad (23)
$$

Both (22) and (23) depend on $\Phi(v)$, which is computed in the Appendix.

In the next lemma, we show that the bounds in (22) and (23) converge as $\gamma \to \infty$.

Lemma 3: Let l be an arbitrary, nonnegative integer.

$$
\begin{aligned}
&\lim_{\gamma\to\infty} 1 + \sum_{k=0}^{l} \frac{1}{k!} \frac{d^k}{dv^k} [e^{2\zeta v}\Phi(-v) - \Phi(v)]\Big|_{v=-1} \\
&\quad = \lim_{\gamma\to\infty} 1 - 4Q\left[\frac{x_a}{4}\sqrt{\frac{\gamma}{\pi}}\right] + \sum_{k=0}^{l} \frac{1}{k!}(2\zeta)^k e^{-2\zeta} \\
&\qquad - \sum_{k=0}^{l} \frac{a^k}{k!} \frac{d^k}{dv^k} \Phi(v)\Big|_{v=-a}.
\end{aligned}
$$

Proof: We rewrite (22) as

$$
\begin{aligned}
&1 - 4Q\left[\frac{x_a}{4}\sqrt{\frac{\gamma}{\pi}}\right] + \sum_{k=0}^{l} \frac{1}{k!}(2\zeta)^k e^{-2\zeta} \\
&\qquad - E[P[N(a\Lambda_0^U) \leq l | \{\Delta\theta_t, 0 \leq t \leq T\}]] \leq 2P[\text{error}].
\end{aligned}
$$

Taking limits of both sides as $\gamma \to \infty$, we have

$$
\begin{aligned}
&P[N(2\zeta) \leq l] + 1 - \lim_{\gamma\to\infty} E[P[N(a\Lambda_0^U) \leq l | \{\Delta\theta_t, 0 \leq t \leq T\}]] \\
&\qquad \leq 2P[\text{error}].
\end{aligned}
$$

The last two terms cancel by an application of the bounded convergence theorem, and the continuity of the Poisson cdf with the mean. The same result follows from the limit of (23) by similar arguments. ■

Since the upper and lower bounds result by replacing the mgf of $1/T\int_0^T \cos\Delta\theta_t\, dt$ by that of $1/T\int_0^T (1 - \Delta\theta_t^2/2)\, dt$, it is of interest to explore the difference between the mgf's of the two random variables. In Fig. 2, we compare the mgf of $1/T\int_0^T \cos\Delta\theta_t\, dt$ via Monte Carlo simulation to the theoretical mgf of $1/T\int_0^T (1 - \Delta\theta_t^2/2)\, dt$ for $\gamma = 30$, which is conservative and well below the range of values of interest in the analysis of the proposed detector. The theoretical mgf of $1/T\int_0^T (1 - \Delta\theta_t^2/2)\, dt$ begins to differ from the mgf of $1/T\int_0^T \cos\Delta\theta_t\, dt$ obtained by Monte Carlo simulation at $v = -6$. As seen in the Appendix, this sharp rise of the mgf is due to a branch point in

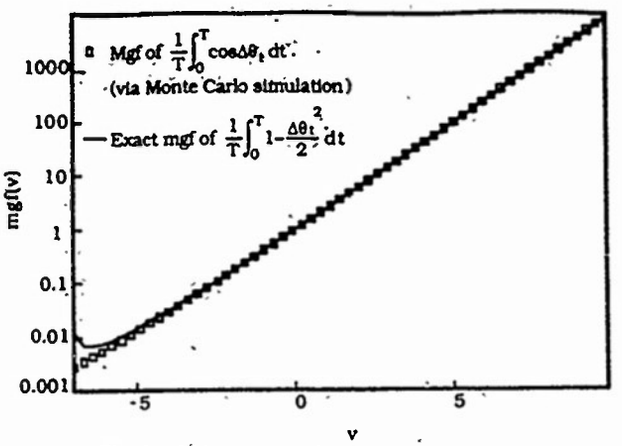

Fig. 2. Moment generating functions ($\gamma \doteq 30$).

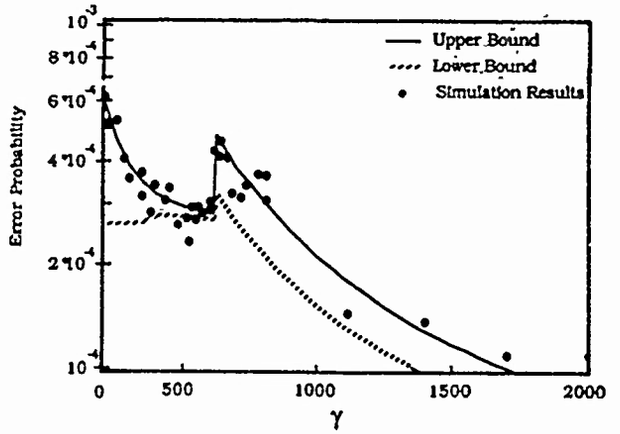

Fig. 3. Error rate bounds and simulation results (6 photons/bit).

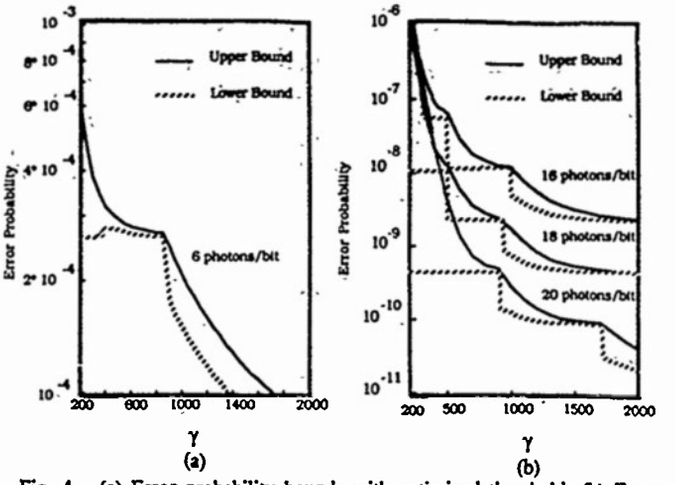

Fig. 4. (a) Error probability bounds with optimized threshold. (b) Error probability bounds with optimized threshold.

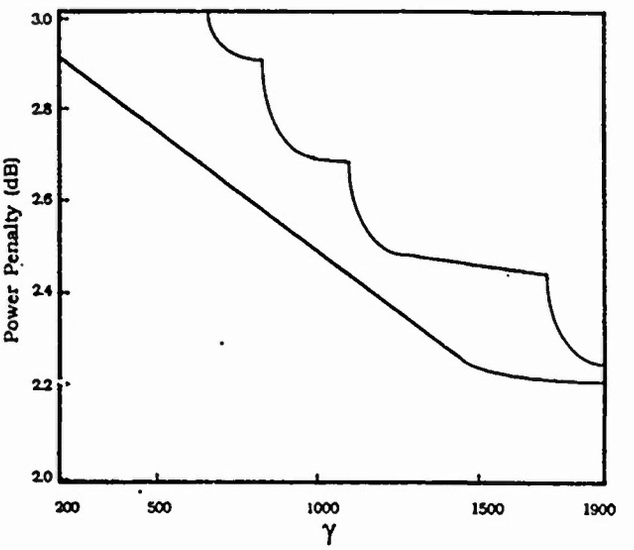

Fig. 5. Bounds on the power penalty of proposed DPSK receiver.

the infinite product expression of the mgf. As γ increases, this branch point occurs for smaller values of v. Fortunately, we are concerned with the region $v \in [-1, 1]$ where we evaluate the mgf for the error probability bounds.

Fig. 3 presents the upper and lower bounds on the error probability of the test in (10) for an optical power of 6 photons per bit. Also shown are results of Monte-Carlo simulations of the hypothesis test. It appears from the simulation results that the upper bound is tighter than the lower bound. This is because the lower bound on $E[P[N(\Lambda_1) \le l | \{\Delta\theta_t, 0 \le t \le T\}]]$ obtained in Lemma 2 is derived by trivially upper bounding $\cos(x)$ by unity for $|x| > x_\epsilon$. In Fig. 3, it can also be seen that the error probability bounds are discontinuous functions of γ. These discontinuities occur for values of (ζ, γ) where the suboptimal threshold, given by the RHS of (10), changes value. Indeed, these discontinuities result from the use of a suboptimal threshold. As the LHS of (10) is integer-valued, it is straightforward to optimize the threshold so as to minimize either the upper or lower bounds. Because of the tightness exhibited by the upper bound, we choose the threshold function which minimizes (23). Additionally, since the process $\{\lambda_t^{(l)}, t \in [0, T]\}$ is close to the mean $E[\lambda_t^{(l)}]$ for moderate γ, the suboptimum threshold function in (10) is equal to the optimized threshold function except in very small intervals in the range of interest of γ. In Fig. 4, we have displayed the lower envelope of the upper bounds corresponding to all integer thresholds, (which is, obviously, an upper bound to the error probability of the test obtained with the optimum threshold) together with the lower bound computed at the threshold that minimizes the upper bound for several values of ζ. That is, Fig. 4 displays (22) and (23) replacing l by the optimized threshold function.

The power penalty is an alternative way to characterize the receiver performance. Fig. 5 shows bounds on the power penalty of the proposed DPSK receiver at 10^{-9} BER. These curves were obtained by recording the values of γ, for fixed optical energy, at which the lower bound (22) and upper bound (23) were equal to 10^{-9}. The optimized threshold was employed for these curves as well. It should be noted that the lower bounding curve in Fig. 5 is a smooth lower bounding envelope to the power penalty data. By comparison, the power penalty for the balanced DPSK receiver as described in [2] is always greater than 3 dB, and attains this value only as $\gamma \to \infty$. As Fig. 5 illustrates, the power penalty for the quantum-limited DPSK receiver is below 3 dB for $\gamma > 700$.

IV. Summary

In this paper, we have analyzed the error probability of an asymptotically quantum-limited direct-detection DPSK receiver. The receiver consists of a delay-and-sum optical preprocessor in tandem with a photoelectric converter and an integrate-and-dump circuit. The output is initially compared to a suboptimal threshold that was derived under the assumption that the conditional rates are constants. We tightly bounded the error probability for arbitrary thresholds by developing upper and lower bounds on the conditional intensities of the photon point process at the photodetector. Prompted by the tightness of the upper bound, we then improved the receiver performance by minimizing the upper bound over all integer threshold

levels. This optimized threshold coincides with the RHS of (10) except in very small intervals in the range of γ. The power penalty bounds were computed using the optimized threshold function, and appear in Fig. 5. While the balanced DPSK receiver analyzed in [2] had a power penalty greater than 3 dB, the receiver presented here had a power penalty less than 3 dB for reasonable values of γ.

APPENDIX

MOMENT GENERATING FUNCTION OF $1/2T \int_0^T \Delta\theta_t^2\, dt$

In this section, we find an expression for the moment generating function of the random variable $1/2T \int_0^T \Delta\theta_t^2\, dt$. A more general problem has been solved previously [12]. Consider the random process $\{z_t \triangleq \int_0^t x_\tau^2 h(t, \tau)\, d\tau,\ t \in \mathbb{R}_+\}$ where $\{x_\sigma,\ \sigma \in \mathbb{R}_+\}$ is a zero mean, wide-sense stationary Gaussian random process with autocorrelation function $R(\tau)$. Then the moment generating function $M_{z_t}(v)$ is given in infinite product form as

$$M_{z_t}(v) = \prod_{i=1}^{\infty} \frac{1}{\sqrt{1-2v\lambda_i}} \tag{A.1}$$

where $\{\lambda_i,\ i = 1, 2 \cdots\}$ are obtained by solving the homogeneous integral equation

$$\lambda_i \phi_i(\sigma) = \int_0^t R(\tau-\sigma) h(t, \tau) \phi_i(\tau)\, d\tau. \tag{A.2}$$

For our particular case, we have

$$t = T$$

$$h(T, \tau) = \frac{1}{2T}$$

and $R(\tau)$ as in (8). By substituting these equations into (A.2), we find for $0 < \sigma < T$

$$\lambda_i \phi_i(\sigma) = \frac{\pi}{\gamma} \int_0^T (T - |\sigma-\tau|) \phi_i(\tau)\, d\tau. \tag{A.3}$$

Similar equations result for other values of σ. Taking the first- and second-partial derivatives of (A.3) with respect to σ we get

$$\lambda_i \dot{\phi}_i(\sigma) = \frac{\pi}{\gamma} \int_0^T \operatorname{sgn}(\tau-\sigma) \phi_i(\tau)\, d\tau \quad 0<\sigma<T \tag{A.4}$$

and

$$\ddot{\phi}_i(\sigma) = -\frac{2\pi}{\gamma\lambda_i} \phi_i(\sigma) \quad 0<\sigma<T. \tag{A.5}$$

Equation (A.5) suggests the general solution

$$\phi_i(\sigma) = A_i \cos \omega_i \sigma + B_i \sin \omega_i \sigma \quad 0<\sigma<T. \tag{A.6}$$

If we substitute (A.6) into (A.3) and (A.4) to solve for the unknowns A_i, ω_i, and B_i, we find that $\{\lambda_i, i = 1, 2 \cdots\}$ are the solutions to

$$\frac{1}{2}\sqrt{\frac{2\pi}{\gamma}\frac{1}{\lambda_i}} \sin \sqrt{\frac{2\pi}{\gamma}\frac{1}{\lambda_i}} = 1 + \cos \sqrt{\frac{2\pi}{\gamma}\frac{1}{\lambda_i}}.$$

Included among these eigenvalues are $\{2\pi/(\gamma[2n+1]^2\pi^2),\ n = 0, 1, 2 \cdots\}$. The remaining portion of the eigenvalues are close to, but not exactly $\{2\pi/(\gamma[2n\pi]^2),\ n = 1, 2 \cdots\}$. Now that the eigenvalues are known, the moment generating function of $1/2T \int_0^T \Delta\theta_t^2\, dt$ may be found by (A.1).

REFERENCES

[1] A. B. Carlson, *Communication Systems, An Introduction to Signals and Noise in Electrical Communication*, 2nd ed. New York: McGraw-Hill, 1975.

[2] J. Salz, "Coherent lightwave communications," *AT&T Tech. J.*, vol. 64, no. 10, pp. 2153–2209, Dec. 1985.

[3] J. E. Mazo, "On binary differential detection for coherent lightwave communication," *AT&T Tech. J.*, vol. 64, no. 10, pp. 2467–2485, Dec. 1985.

[4] D. L. Snyder, *Random Point Processes*. New York: Wiley, 1975.

[5] P. Bremaud, *Point Processes and Queues: Martingale Dynamics*. New York: Springer-Verlag, 1981.

[6] M. Françon, *Optical Interferometry*. New York: Academic, 1966.

[7] G. L. Abbas, V. W. S. Chan, and T. K. Yee, "A dual-detector optical heterodyne receiver for local oscillator noise suppression," *IEEE J. Lightwave Technol.*, vol. LT-3, pp. 1110–1122, Oct. 1985.

[8] E. E. Basch, Ed., *Optical-Fiber Transmission*. Indianapolis, IN: Howard W. Sams & Co., Inc., 1987.

[9] D. Slepian, "First passage time for a particular Gaussian process," *Ann. Math. Stat.*, vol. 32, pp. 610–612, 1961.

[10] N. Wiener, A. Siegel, B. Rankin, and W. T. Martin, *Differential Space, Quantum Systems, and Prediction*. Cambridge, MA: M.I.T. Press, pp. 78–79.

[11] D. A. Darling and A. J. F. Siegert, "The first passage problem for a continuous Markov process," *Ann. Math. Stat.*, vol. 24, pp. 624–639, 1953.

[12] J. B. Thomas and E. Wong, "The distribution of quadratic functionals," *Quarter. Appl. Math.*, vol. 19, no. 2, pp. 150–153, July, 1961.

★

David Brady (S'84) was born on June 23, 1959, in West Chester, PA. He received the B.S. degree in electrical engineering from Carnegie-Mellon University (with university honors) in 1982, and the M.S.E.E. degree from the California Institute of Technology in 1983. He is currently working towards the Ph.D. degree in electrical engineering at Princeton University, Princeton, NJ.

During 1983–1984, he was a member of Technical Staff at Bell Laboratories, Holmdel, NJ, and from 1984 until 1985, he was a member of Technical Staff at Bell Communications Research, Inc., West Long Branch, NJ, where he worked on the characterization of phase jitter in asynchronous DS-X networks. His research interests include optical communications, multiuser communication systems, and information theory.

Mr. Brady is a member of Tau Beta Pi and Eta Kappa Nu.

★

Sergio Verdú (S'80–M'84–SM'88) received the Ph.D. degree in electrical engineering from the University of Illinois at Urbana-Champaign in 1984.

Upon completion of his doctorate he joined the Faculty of Princeton University, Princeton, NJ, where he is an Assistant Professor of Electrical Engineering. His current research interests are in the areas of multiuser communication and information theory, and statistical signal processing.

Dr. Verdú is a recipient of the National University Prize of Spain, the Rheinstein Outstanding Junior Faculty Award of the School of Engineering and Applied Science in Princeton University, and the NSF Presidential Young Investigator Award. He is currently serving as Associate Editor of the IEEE TRANSACTIONS ON AUTOMATIC CONTROL.

Multiple-Access Channels with Memory with and without Frame Synchronism

Sergio Verdú

Reprinted from
IEEE TRANSACTIONS ON INFORMATION THEORY
Vol. 35. No. 3, May 1989

Multiple-Access Channels with Memory with and without Frame Synchronism

SERGIO VERDÚ, SENIOR MEMBER, IEEE

Abstract—The capacity region of frame-synchronous and asynchronous discrete two-user multiple-access channels with finite memory is obtained. Frame synchronism refers to the ability of the transmitters to send their codewords in unison. The absence of frame synchronism in memoryless multiple-access channels is known to result in the removal of the convex hull operation from the expression of the capacity region. We show that when the channel has memory, frame asynchronism rules out nonstationary inputs to achieve any point in the capacity region, thereby allowing only coding strategies that involve cooperation in the frequency domain but not in the time domain. This restriction drastically reduces the capacity region of some multiple-access channels with memory, and in particular the total capacity of the channel, which is invariant to frame asynchronism for memoryless channels.

I. Introduction

THE CENTRAL result in multiuser information theory states that the capacity region of a two-user discrete multiple-access channel is equal to the convex closure of the set of rate pairs (R_1, R_2) satisfying

$$\begin{aligned} 0 \le R_1 &\le I(X; Z|Y), \\ 0 \le R_2 &\le I(Y; Z|X), \\ R_1 + R_2 &\le I(X, Y; Z) \end{aligned} \tag{1}$$

for some independent input distributions X and Y and output distribution Z. This result was obtained by Ahlswede [1] in 1971 under the key assumptions that the channel is memoryless and frame (or block)-synchronous, i.e., the beginnings of the codewords sent by the transmitters coincide. In the absence of frame synchronism, an unpredictable offset exists between the epochs at which the codewords of each user are received at the decoder. Even though the receiver can easily acquire timing synchronism with each user and hence know the value of the offset prior to decoding, the transmitters must encode their messages without knowing the offset between their codewords. Assuming that the offset can be guaranteed to be negligible with respect to the codeword length (e.g., if an upper bound on the offset is known by the transmitters), Cover *et al.* [5] and Narayan and Snyder [15] proved that the capacity region and the cutoff rate region, respectively, of the discrete memoryless multiple-access channel are the same as in the frame-synchronous case. Poltyrev [18] and, independently, Hui and Humblet [12] have shown that if no information on the actual value of the offset is available to the transmitters (i.e., if the channel is "completely" frame-synchronous), then the capacity region of the discrete memoryless multiple-access channel is as stated above but without the convex hull operation. This implies that in most memoryless channels of interest, frame asynchronism does not change the capacity region, the most well-known exceptions being, perhaps, the Massey–Mathys collision channel without feedback [13] and the counterexamples in [6, p. 287] and [3]. Furthermore, the Poltyrev–Hui–Humblet result implies that the maximum achievable rate sum $(R_1 + R_2)$ (or total capacity) of the memoryless multiple-access channel is never decreased by the lack of frame synchronism, because the rate sum of any convex combination of rate pairs is equal to the convex combination of the respective rate sums. As we shall see, these conclusions are no longer true when the multiple-access channel has memory.

Even though the study of the capacity of single-user channels with memory has occupied a prominent position in the development of the Shannon theory, multiple-access channels with memory have received scant attention in the literature (see van der Meulen [14, open problem 12]). Aside from their inherent conceptual and practical interest, multiple-access channels with memory play a key role in the modeling of symbol-asynchronous channels.[1] These are continuous-time channels where each codeword symbol modulates a finite-duration signal waveform and the transmitters do not cooperate so that the symbol epochs are aligned at the receiver. Since each symbol overlaps with two consecutive symbols transmitted by the other user, the equivalent discrete-time multiple-access channels required to model symbol-asynchronous channels have memory [21]. Therefore, the study of multiple-access channels where the transmitters are completely asynchronous leads to frame-asynchronous discrete-time multiple-access channels with memory.

Manuscript received November 23, 1987; revised August 25, 1988. This work was supported in part by the Office of Naval Research under contract N00014-87-K-0054. The material in this paper was partially presented at the 1988 IEEE International Symposium on Information Theory, Kobe, Japan, June 1988.

The author is with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544.

IEEE Log Number 8928185.

[1] One more type of channel "asynchronism" is that which allows deletions and insertions of symbols at locations unknown to the decoder. This has been studied by Dobrushin [7] and by Ahlswede and Gacs [2] in the context of single-user and multiple-access memoryless channels, respectively.

0018-9448/89/0500-0605$01.00

The multiple-access channel with memory studied in this paper has finite input alphabets A_1 and A_2 and finite output alphabet B. Except for the general converse theorem for synchronous channels proved in Section II, our results are obtained under the assumption that the multiple-access channel is stationary and has finite (input) memory, in the sense that each channel output depends on up to m consecutive inputs of each user, and the outputs are conditionally independent given the inputs.[2] The multiple-access channel with finite memory encompasses many cases of practical interest such as the symbol-asynchronous channel and channels with finite-length intersymbol interference; its capacity has been solved in the single-user case in the works of Tsaregradsky [20], Feinstein [8], and Wolfowitz [22].

As usual when dealing with sources or channels with memory, the capacity region of the multiple-access channel with memory does not admit single-letter characterizations and, rather, is given in terms of a limit of regions. This fact does not curtail the applicability or interest of these results because those limits are computable, as we show in several examples where they result in explicit closed-form expressions. Moreover, we provide a theorem (which generalizes Wolfowitz's result [22, theorem 5.5.1] on the speed of convergence of single-user capacity) that allows the computation of the capacity region of the channel with memory up to any desired degree of approximation via the computation of achievable regions for memoryless channels.

As in the case of the memoryless multiple-access channel, the frame-synchronous channel with memory is shown to satisfy the time-sharing principle, i.e., its capacity region is convex. As a form of cooperation in the time domain, time-sharing requires nonstationary input distributions. Note that, while stationary inputs always achieve capacity in time-invariant single-user channels, there are time-invariant multiple-access channels (e.g., the aforementioned channels whose capacity region is decreased without frame-synchronism) that require nonstationary inputs to achieve all points in the capacity region. In this paper we show that only stationary inputs are allowed for frame-asynchronous multiple-access channels with memory. Hence cooperation between the users is beneficial in the frequency domain (dependent inputs are necessary to achieve capacity because the channel has memory) but not in the time domain due to the lack of a common time reference. The opposite situation is encountered in the frame-synchronous memoryless multiple-access channel, where it is enough to restrict attention to independent input sequences and time-sharing (hence nonstationary) inputs may be required to achieve capacity. In the light of our results, the Poltyrev–Hui–Humblet result for memoryless channels is a consequence of the nonstationarity of time-sharing strategies.

[2]The results and proof techniques of this paper easily generalize to the case when the outputs are conditionally m-dependent given the inputs, i.e., when all pairs of subsets of random variables whose indices differ by more than m are independent.

The actual impact that the lack of frame synchronism (i.e., the restriction to stationary inputs) has on the capacity region of the multiple-access channel with memory is quite diverse. On one hand, there are many frame-synchronous channels (e.g., the symbol-asynchronous multiple-access channel considered in [21]) whose capacity regions are achieved by stationary inputs, and therefore, they do not decrease if the users are not guaranteed to transmit their codewords in unison. On the other hand, we show in this paper the existence of channels with memory where not only the capacity region but the total capacity is drastically reduced by the lack of frame synchronism.

II. Frame-Synchronous Capacity Region

We give first a general converse coding theorem for the discrete frame-synchronous multiple-access channel that puts no restrictions on its transition probabilities.

Theorem 1: The capacity region of the discrete frame-synchronous multiple-access channel satisfies[3]

$$\mathcal{C} \subset \text{closure}\left\{ \liminf_{n\to\infty} \frac{1}{n} C_n \right\} \tag{2}$$

where

$$C_n = \bigcup_{X^n Y^n} \Big\{ (R_1, R_2): 0 \le R_1 \le I(X^n; Z^n | Y^n)$$
$$0 \le R_2 \le I(Y^n; Z^n | X^n)$$
$$R_1 + R_2 \le I(X^n, Y^n; Z^n) \Big\} \tag{3}$$

and the union is over independent n-dimensional input distributions. Note that the convex closure of C_n is the capacity region of the discrete memoryless multiple-access channel whose input and output alphabets are A_1^n, A_2^n, and B^n, respectively, and whose transition probabilities are $P_{Z^n|X^n,Y^n}$.

Proof: We need to show that, for all $0 < \epsilon < 1$, every ϵ-achievable rate pair (R_1, R_2) belongs to the right side of (2). If (R_1, R_2) is ϵ-achievable, then for all $\gamma > 0$ and for all sufficiently large n there exists an (n, M_1, M_2, ϵ) code (i.e., a code with block length n, M_i codewords for user i, and average probability that both messages are correctly decoded greater or equal than $1-\epsilon$) such that

$$\frac{\log M_i}{n} \ge R_i - \gamma, \qquad i = 1, 2. \tag{4}$$

Fix one such code and let S_1 and S_2 denote independent random variables uniformly distributed on $\{1, \cdots, M_1\}$, and $\{1, \cdots, M_2\}$. The message transmitted by user i is a realization of S_i. Let Z^n denote the output of the channel, when S_1 and S_2 are transmitted using the above (n, M_1, M_2, ϵ) code, and let $(\hat{S}_1, \hat{S}_2)$ be the messages se-

[3]All the logarithms, exponentials, entropies and mutual informations in this paper have a common arbitrary base.

lected by the decoder. The Fano inequality states that

$$H(S_1, S_2|Z^n) \le \log 2 + P\left[(S_1, S_2) \ne (\hat{S}_1, \hat{S}_2)\right] \log M_1 M_2 \quad (5a)$$

$$H(S_1|Z^n) \le \log 2 + P\left[S_1 \ne \hat{S}_1\right] \log M_1 \quad (5b)$$

$$H(S_2|Z^n) \le \log 2 + P\left[S_2 \ne \hat{S}_2\right] \log M_2. \quad (5c)$$

Since the average probability of error of the code does not exceed ϵ, the probabilities in the upper bounds of (5) can be replaced by ϵ, and because S_1 and S_2 are uniformly distributed, we can write

$$I(S_1; Z^n) \ge (1-\epsilon)\log M_1 - \log 2 \quad (6a)$$

$$I(S_2; Z^n) \ge (1-\epsilon)\log M_2 - \log 2 \quad (6b)$$

$$I(S_1, S_2; Z^n) \ge (1-\epsilon)\log M_1 M_2 - \log 2. \quad (6c)$$

If f_i: $\{1, \cdots, M_i\} \to A_i^n$ denotes the code book of user i, then since S_1 and S_2 are independent,

$$\begin{aligned} I(f_1(S_1); Z^n|f_2(S_2)) &= I(f_1(S_1); f_2(S_2)) \\ &\quad + I(f_1(S_1); Z^n|f_2(S_2)) \\ &= I(f_1(S_1); Z^n, f_2(S_2)) \\ &\ge I(f_1(S_1); Z^n) \\ &\ge I(S_1; Z^n) \end{aligned} \quad (7)$$

where the last inequality follows from the data-processing lemma. In a similar way, we obtain

$$I(f_2(S_2); Z^n|f_1(S_1)) \ge I(S_2; Z^n) \quad (8a)$$

$$I(f_1(S_1), f_2(S_2); Z^n) \ge I(S_1, S_2; Z^n). \quad (8b)$$

Now, putting (4) together with (6)–(8), we get

$$(1-\epsilon)(R_1 - \gamma) - \frac{\log 2}{n} \le \frac{1}{n} I(f_1(S_1); Z^n|f_2(S_2))$$

$$(1-\epsilon)(R_2 - \gamma) - \frac{\log 2}{n} \le \frac{1}{n} I(f_2(S_2); Z^n|f_1(S_1))$$

$$(1-\epsilon)(R_1 + R_2 - 2\gamma) - \frac{\log 2}{n} \le \frac{1}{n} I(f_1(S_1), f_2(S_2); Z^n)$$

which implies that

$$(1-\epsilon)(R_1 - \gamma, R_2 - \gamma) \div \left(\frac{\log 2}{n}, \frac{\log 2}{n}\right) \in \frac{1}{n} C_n$$

for all sufficiently large n, and consequently,

$$(1-\epsilon)(R_1 - 2\gamma, R_2 - 2\gamma) \in \frac{1}{n} C_n$$

or all sufficient large n, or in other words

$$(1-\epsilon)(R_1 - 2\gamma, R_2 - 2\gamma) \in \liminf_{n \to \infty} \frac{1}{n} C_n. \quad (9)$$

However, since ϵ and γ are arbitrarily small, (9) implies hat (R_1, R_2) must be the limit of a sequence of points belonging to $\liminf_{n\to\infty}(1/n)C_n$, and therefore it belongs o the right side of (2) (as was to be shown).

As in the case of the single-user channel with memory, no universal direct coding theorem is known to hold with the same generality as Theorem 1. We will henceforth focus our attention on the following class of multiple-access channels with memory.

Definition: A multiple-access channel with finite memory m is one whose channel transition probability satisfies

$$\begin{aligned} &p_{Z^n|X^n,Y^n}(w_1, \cdots, w_n|a_1, \cdots, a_n, b_1, \cdots, b_n) \\ &\quad = p_{Z_1, \cdots, Z_{m-1}|X_1, \cdots, X_{m-1}, Y_1, \cdots, Y_{m-1}} \\ &\qquad \cdot (w_1, \cdots, w_{m-1}|a_1, \cdots, a_{m-1}, b_1, \cdots, b_{m-1}) \\ &\qquad \cdot \prod_{i=m}^{n} p_c(w_i|a_{i-m+1}, \cdots, a_i, b_{i-m+1}, \cdots, b_i) \end{aligned} \quad (10)$$

for all $n > 0$.

This implies that the outputs $Z_m, \cdots, Z_n$ are conditionally independent given the inputs, and each of them depends on m consecutive inputs of each user, thus encompassing intersymbol interference of finite duration. This definition allows us to handle the boundary outputs $Z_1, \cdots, Z_{m-1}$ (which depend on fewer than m input symbols from each user) in any arbitrary way, and it is therefore preferable to the single-user definition of [8] and [22] where the boundary outputs are not available to the decoder. As we shall see, the capacity region of the multiple-access channel with memory depends only on the transition probability p_c and not on the conditional distribution of the first $m-1$ outputs.

In what follows it is convenient to refer to a memoryless multiple-access channel derived from the channel with memory in the following way.

Definition: Let $l \ge m$. The l-block multiple-access channel derived from a multiple-access channel with finite memory m is a memoryless channel characterized by input alphabets A_1^l and A_2^l, output alphabet B^{l-m+1}, and transition probability

$$\begin{aligned} &p((w_m, \cdots, w_l)|(a_1, \cdots, a_l), (b_1, \cdots, b_l)) \\ &\quad = \prod_{i=m}^{l} p_c(w_i|a_{i-m+1}, \cdots, a_i, b_{i-m+1}, \cdots, b_i). \end{aligned} \quad (11)$$

It follows from (1) that the capacity region of the l-block memoryless multiple-access channel is equal to the convex closure of

$$\begin{aligned} Q_l = \bigcup_{X^l, Y^l} \Big\{ (R_1, R_2): 0 \le R_1 \le I(X^l; Z_m^l|Y^l) \\ 0 \le R_2 \le I(Y^l; Z_m^l|X^l) \\ R_1 + R_2 \le I(X^l, Y^l; Z_m^l) \Big\} \end{aligned} \quad (12)$$

where the union is over independent distributions on the sets A_1^l and A_2^l, respectively, and $Z_m^l = (Z_m, \cdots, Z_l)$. The direct coding theorem for the multiple-access channel with finite memory gives the following achievable region as a function of the achievable region in (12) for the l-block multiple-access channel.

Theorem 2: The capacity region of the frame-synchronous multiple-access channel with finite memory m satis-

ies

$$\supset \text{closure}\left(\bigcup_{l\geq m}\frac{1}{l}Q_l\right). \tag{13}$$

Proof: We need to show that, for all $l \geq m$ and $(R_1, R_2) \in (1/l)Q_l$ and for every $0 < \epsilon < 1$ and $\gamma > 0$, there exist (n, M_1, M_2, ϵ) codes for all sufficiently large n such nat

$$\frac{\log M_i}{n} \geq R_i - \gamma, \qquad i = 1,2.$$

Then $\bigcup_{l\geq m}(1/l)Q_l$ will be an achievable region and so will its closure since the capacity region is a closed set.

First, we will fix l and show the existence of said codes for sufficiently large multiples of l: $n = kl$. Since Q_l is an achievable region of the l-block memoryless multiple-access channel. if $(R_1', R_2') \in Q_l$, then for every $\gamma_l > 0$ and all k sufficiently large, there exist (k, M_1, M_2, ϵ) codes for the l-block channel such that

$$\frac{\log M_i}{k} \geq R_i' - \gamma_l, \qquad i = 1,2. \tag{14}$$

Now. we fix one such code and view the symbols in each of its codewords sequentially. In this way, we have a code for the multiple-access channel with memory with block length kl and M_i codewords for user i. Its probability of error is not greater than ϵ because if we were to constrain the decoder not to use the outputs

$$Z_1, \cdots, Z_{m-1}, Z_{l+1}, \cdots, Z_{l+m-1}, \cdots, Z_{(k-1)l+1}, \cdots, Z_{(k-1)l+m-1},$$

then the situation would be entirely equivalent to decoding in the l-block memoryless channel where there is no interference between the l-blocks. Clearly, if those outputs are not discarded. the probability of error cannot increase. Now letting

$$R_1', R_2') = l(R_1, R_2) \in Q_l \qquad \gamma_l = \frac{\gamma}{2}$$

14) results in

$$\frac{\log M_i}{kl} \geq R_i - \frac{\gamma}{2} \tag{15}$$

as we wanted to show. However. this only proves the existence of reliable codes with the desired rates for block lengths that are multiples of l. To find codes whose block length is $n = kl + t$, $t = 1, \cdots, l-1$, we append t arbitrary input symbols to each of the codewords of the foregoing $kl, M_1, M_2, \epsilon)$ codes, and let the decoder discard the last t outputs. Then, it is clear that the probability of error remains unchanged and the rates of the new codes satisfy via (15))

$$\frac{\log M_i}{kl+t} \geq R_i - \frac{t}{kl+t}R_i - \frac{kl}{kl+t}\frac{\gamma}{2}, \qquad i = 1,2. \tag{16}$$

For sufficiently large k, however, $tR_i/(lk+t) < \gamma/2$ in which case the right side of (16) can be further lower-bounded by $R_i - \gamma$. Thus even though the new code has lower rates than the original code with block length kl, the decrease is inappreciable for large k.

The following result proves that the inner and outer bounds shown in Theorems 1 and 2 coincide.

Theorem 3: The capacity region of the frame-synchronous multiple-access channel with finite memory m is given by

$$C = \text{closure}\left(\bigcup_{n\geq m}\frac{1}{n}Q_n\right) = \text{closure}\left(\liminf_{n\to\infty}\frac{1}{n}Q_n\right)$$
$$= \text{closure}\left(\limsup_{n\to\infty}\frac{1}{n}C_n\right) = \text{closure}\left(\liminf_{n\to\infty}\frac{1}{n}C_n\right). \tag{17}$$

Proof: The essence of the proof is the following inequality which holds for all $l \geq m$:

$$I(X^l, Y^l; Z^l) = I(X^l, Y^l; Z_m^l) + I(X^l, Y^l; Z^{m-1}|Z_m^l)$$
$$\leq I(X^l, Y^l; Z_m^l) + (m-1)\log|B| \tag{18a}$$

and simularly,

$$I(X^l; Z^l|Y^l) \leq I(X^l; Z_m^l|Y^l) + (m-1)\log|B| \tag{18b}$$

$$I(Y^l; Z^l|X^l) \leq I(Y^l; Z_m^l|X^l) + (m-1)\log|B| \tag{18c}$$

which imply that

$$C_l \subset Q_l + (m-1)\log|B|U \tag{19}$$

where U is the unit square $\{(x_1, x_2): 0 \leq x_1 \leq 1, 0 \leq x_2 \leq 1\}$. We now have the following chain of inclusions:

$$\text{closure}\left(\limsup_{n\to\infty}\frac{1}{n}C_n\right) \subset \text{closure}\left(\limsup_{n\to\infty}\frac{1}{n}Q_n\right)$$
$$\subset \text{closure}\left(\bigcup_{n\geq m}\frac{1}{n}Q_n\right)$$
$$\subset C$$
$$\subset \text{closure}\left(\liminf_{n\to\infty}\frac{1}{n}C_n\right)$$
$$\subset \text{closure}\left(\liminf_{n\to\infty}\frac{1}{n}Q_n\right) \tag{20}$$

where the first and last inclusions follow easily from (19) and the third and fourth inclusions are Theorems 2 and 1, respectively. Finally, since the liminf is a subset of the limsup all the inclusions in (20) are in fact equalities.

The closure operation in (17) is indeed necessary because even if $\lim_{n\to\infty}(Q_n/n)$ exists, it may not be a closed set (e.g., if the first $m-1$ boundary outputs are independent of the inputs). At first sight it may seem surprising that the capacity region of Theorem 3 does not involve an explicit convex hull operation, especially in light of the fact that the particular case of the frame-synchronous memoryless multiple-access channel is known to require the convex hull operation. In fact the capacity region of Theorem 3 is already convex because it is given as a limit of achievable regions for n-block channels whose input distributions are allowed to time-share among several distributions as a

result of the assumption that the users are frame-synchronous. This is formalized in the following result.

Corollary (Time-Sharing Principle): The capacity region of the frame-synchronous multiple-access channel with finite memory is a convex set.

Proof: It follows from Theorem 3 that the capacity region is independent of the conditional probability of the $(m-1)$ boundary outputs; therefore, we can prove the corollary for any arbitrary choice of this probability; in particular, we shall assume that $Z_1, \cdots, Z_{m-1}$ are independent of the inputs. Then we can follow the same approach as in the proof of the time-sharing principle for memoryless channels [6] which juxtaposes two codes. If we impose the restriction that the decoder must discard the leading $(m-1)$ output symbols of each of the two blocks, then the decoding of the new code is decoupled and equivalent to the case when the codewords are sent individually. Therefore, the error probability of the new code is better than the sum of the probabilities of error of the two component codes, the rate pair is a convex combination of the rate pairs of both codes, and the proof proceeds as in the memoryless case [6, p. 272].

Theorem 3 can easily be generalized in several directions. For example, the proof of both the converse and the direct theorems remain essentially unchanged for continuous-alphabet channels with input constraints. Another generalization which is of interest in the symbol-asynchronous channel [21] is that of a compound multiple-access channel where the transmitters only know that the channel belongs to an uncertainty set Γ (cf. [6, p. 288] for the corresponding memoryless result). In that problem, the proof of the direct theorem requires very little modification since the construction of codebooks therein is independent of the channel, and the proof of the converse only needs to take care of the fact that a good code must be so for any possible channel in the uncertainty set. Then Theorem 3 holds by replacing C_n by

$$C_n' = \bigcup_{X^n Y^n} \bigcap_{\omega \in \Gamma} \{(R_1, R_2): 0 \le R_1 \le I(X^n; Z^n(\omega)|Y^n) \quad 0 \le R_2 \le I(Y^n; Z^n(\omega)|X^n) \quad R_1 + R_2 \le I(X^n, Y^n; Z^n(\omega))\}$$

where $Z^n(\omega)$ is connected to X^n and Y^n through channel $\omega \in \Gamma$.

For the purposes of illustration we will show several examples where the limits of Theorem 3 are explicitly computable. However, in cases without much structure an alternative to the analytical computation of those limits is their numerical approximation. This can be done using the following theorem, which allows the computation of the capacity region as accurately as desired via the computation of achievable regions for memoryless channels. Theorem 4 is a generalization of the single-user result obtained by Wolfowitz [22, theorem 5.5.1].[4]

[4]Added in proof: Theorem 4 gives a solution to Problem 1 in [23].

Theorem 4: The capacity region of the frame-synchronous multiple-access channel with finite memory satisfies for every $l \ge m$

$$\text{closure}\left(\text{convex}\frac{1}{l}Q_l\right) \subset C \subset \text{closure}\left(\text{convex}\frac{1}{l}Q_l\right) + \frac{m-1}{l}\log|B|U \quad (21)$$

where U is the unit square $\{(x_1, x_2): 0 \le x_1 \le 1, 0 \le x_2 \le 1\}$.

Proof: The inner bound is a consequence of Theorem 2 and the corollary to Theorem 3. To show the upper bound, fix $l \ge m$ and notice that for any $n = kl$, X^n, and Y^n,

$$\begin{aligned}
I(X^n, Y^n; Z^n) &= I\left(X^n, Y^n; Z_m^l Z_{m+l}^{2l} \cdots Z_{m+(k-1)l}^{kl}\right) \\
&\quad + I\left(X^n, Y^n; Z_1^{m-1} Z_{l+1}^{l+m+1} \times \cdots Z_{(k-1)l+1}^{(k-1)l+m-1} | Z_m^l \cdots Z_{m+(k-1)l}^{kl}\right) \\
&\le k(m-1)\log|B| \\
&\quad + I\left(X^n, Y^n; Z_m^l \cdots Z_{m+(k-1)l}^{kl}\right) \\
&= k(m-1)\log|B| + H\left(Z_m^l \cdots Z_{m+(k-1)l}^{kl}\right) \\
&\quad - H\left(Z_m^l \cdots Z_{(k-1)l+m}^{kl} | X^n, Y^n\right) \\
&\le k(m-1)\log|B| + \sum_{j=0}^{k-1} H\left(Z_{m+jl}^{l+jl}\right) \\
&\quad - H\left(Z_m^l \cdots Z_{(k-1)l+m}^{kl} | X^n, Y^n\right) \\
&= k(m-1)\log|B| \\
&\quad + \sum_{j=0}^{k-1} \left\{H\left(Z_{m+jl}^{l+jl}\right) - H\left(Z_{m+jl}^{l+jl} | X_{1+jl}^{l+jl}, Y_{1+jl}^{l+jl}\right)\right\} \\
&= k(m-1)\log|B| \\
&\quad + \sum_{j=0}^{k-1} I\left(X_{1+jl}^{l+jl}, Y_{1+jl}^{l+jl}; Z_{m+jl}^{l+jl}\right) \qquad (22)
\end{aligned}$$

where the next-to-last identity follows from the definition of the channel with finite memory. Similarly, we can upper-bound

$$I(X^n; Z^n|Y^n) \le k(m-1)\log|B| + \sum_{j=0}^{k-1} I\left(X_{1+jl}^{l+jl}; Z_{m+jl}^{l+jl} | Y_{1+jl}^{l+jl}\right) \quad (23a)$$

$$I(Y^n; Z^n|X^n) \le k(m-1)\log|B| + \sum_{j=0}^{k-1} I\left(Y_{1+jl}^{l+jl}; Z_{m+jl}^{l+jl} | X_{1+jl}^{l+jl}\right). \quad (23b)$$

However, we saw in the proof of Theorem 1 that if (R_1, R_2) is ϵ-achievable, then for all $\gamma > 0$ and for all sufficiently large n, there exist (n, M_1, M_2, ϵ) codes such that

$$\frac{\log M_i}{n} \ge R_i - \gamma, \qquad i = 1, 2 \quad (24)$$

and for some input distributions X^n and Y^n,

$$(1-\epsilon)\log M_1 - \log 2 \le I(X^n; Z^n|Y^n) \quad (25a)$$

$$(1-\epsilon)\log M_2 - \log 2 \le I(Y^n; Z^n|X^n) \quad (25b)$$

$$(1-\epsilon)\log M_1M_2 - \log 2 \le I(X^n, Y^n; Z^n). \quad (25c)$$

Now, combining (22)–(25), we obtain for all sufficiently large k

$$\frac{1}{k}\left[(1-\epsilon)\log M_1 - \log 2\right] - (m-1)\log|B| \le \frac{1}{k}\sum_{j=0}^{k-1} I\left(X_{1+jl}^{l+jl}; Z_{m+jl}^{l+jl}|Y_{1+jl}^{l+jl}\right)$$

$$\frac{1}{k}\left[(1-\epsilon)\log M_2 - \log 2\right] - (m-1)\log|B| \le \frac{1}{k}\sum_{j=0}^{k-1} I\left(Y_{1+jl}^{l+jl}; Z_{m+jl}^{l+jl}|X_{1+jl}^{l+jl}\right)$$

$$\frac{1}{k}\left[(1-\epsilon)\log M_1M_2 - \log 2\right] - (m-1)\log|B| \le \frac{1}{k}\sum_{j=0}^{k-1} I\left(X_{1+jl}^{l+jl}, Y_{1+jl}^{l+jl}; Z_{m+jl}^{l+jl}\right)$$

which implies that

$$\frac{1}{k}(1-\epsilon)(\log M_1, \log M_2) \in \text{convex}\{Q_l\} + \left(\frac{\log 2}{k} + (m-1)\log|B|\right)U$$

for all sufficiently large k. This together with (24) implies that any ϵ-achievable pair (R_1, R_2) satisfies

$$(1-\epsilon)(R_1, R_2) \in \text{closure}\left(\text{convex}\frac{1}{l}Q_l\right) + \frac{m-1}{l}\log|B|U. \quad (26)$$

Thus if (R_1, R_2) is an achievable pair, then it must belong to the closed set in the right side of (26).

The following examples serve to illustrate the analytical evaluation of the capacity region of the frame-synchronous multiple-access channel with memory. In Section III, we derive the capacity of these channels in the absence of synchronism.

Example 1: Consider the following multiple-access channel with finite ($m = 2$) memory which is a simple discrete-time noiseless model of two-user duobinary transmission: $A_1 = A_2 = \{0,1\}$, $B = \{0,1,2,3,4\}$

$$z_i = x_i + x_{i-1} + y_i + y_{i-1} \quad (27)$$

where, according to Theorem 3, it is not necessary to specify the initial conditions as far as computing the capacity region is concerned. To evaluate C, first we compute the mutual informations in the definition of Q_n (12). Since the outputs are deterministic given the inputs, we have

$$I(X_1,\cdots,X_n,Y_1,\cdots,Y_n;Z_2,\cdots,Z_n) = H(Z_2,\cdots,Z_n) \quad (28a)$$

$$I(X_1,\cdots,X_n;Z_2,\cdots,Z_n|Y_1,\cdots,Y_n) = H(Z_2,\cdots,Z_n|Y_1,\cdots,Y_n) \quad (28b)$$

$$I(Y_1,\cdots,Y_n;Z_2,\cdots,Z_n|X_1,\cdots,X_n) = H(Z_2,\cdots,Z_n|X_1,\cdots,X_n). \quad (28c)$$

Moreover, the properties of conditional entropy result in

$$H(Z_2,\cdots,Z_n|X_1+Y_1) \le H(Z_2,\cdots,Z_n) \le H(Z_2,\cdots,Z_n|X_1+Y_1) + H(X_1+Y_1) \quad (29)$$

and

$$\begin{aligned} H(Z_2,\cdots,Z_n|X_1+Y_1) &= H(Z_2|X_1+Y_1) + H(Z_3|X_1+Y_1,Z_2) + \cdots + H(Z_n|X_1+Y_1,Z_2,\cdots,Z_{n-1}) \\ &= H(X_2+Y_2|X_1+Y_1) + H(X_3+Y_3|X_1+Y_1,X_2+Y_2) + \cdots + H(Z_n|X_1+Y_1,\cdots,X_{n-1}+Y_{n-1}) \\ &= H(X_1+Y_1,\cdots,X_n+Y_n) - H(X_1+Y_1). \end{aligned} \quad (30)$$

Also, using the definition of the channel and the fact that $(X_1,\cdots,X_n)$ and $(Y_1,\cdots,Y_n)$ are mutually independent, we can write

$$H(Z_2,\cdots,Z_n|Y_1,\cdots,Y_n) = H(X_1+X_2,\cdots,X_{n-1}+X_n|Y_1,\cdots,Y_n) = H(X_1+X_2,\cdots,X_{n-1}+X_n) \quad (31a)$$

and, similarly,

$$H(Z_2,\cdots,Z_n|X_1,\cdots,X_n) = H(Y_1+Y_2,\cdots,Y_{n-1}+Y_n). \quad (31b)$$

Now, putting together (28)–(31), we obtain

$$\begin{aligned} C &= \text{closure}\left(\liminf_{n\to\infty}\frac{1}{n}Q_n\right) \\ &= \text{closure}\left(\liminf_{n\to\infty}\bigcup_{X^nY^n}\left\{(R_1,R_2):\right.\right. \\ &\quad 0 \le R_1 \le \frac{1}{n}H(X_1+X_2,\cdots,X_{n-1}+X_n) \\ &\quad 0 \le R_2 \le \frac{1}{n}H(Y_1+Y_2,\cdots,Y_{n-1}+Y_n) \\ &\quad \left.\left. R_1+R_2 \le \frac{1}{n}H(X_1+Y_1,\cdots,X_n+Y_n)\right\}\right) \\ &= \text{closure}\left(\liminf_{n\to\infty}\bigcup_{X^nY^n}\left\{(R_1,R_2):\right.\right. \\ &\quad 0 \le R_1 \le \frac{1}{n}H(X_1,\cdots,X_n) \\ &\quad 0 \le R_2 \le \frac{1}{n}H(Y_1,\cdots,Y_n) \\ &\quad \left.\left. R_1+R_2 \le \frac{1}{n}H(X_1+Y_1,\cdots,X_n+Y_n)\right\}\right) \end{aligned} \quad (32)$$

and since each of the three entropies in (32) is maximized by independent equiprobable inputs ($\max_{XY} H(X+Y)$ over independent binary X and Y is equal to 1.5 bit and is achieved by equiprobable distributions), the right side of (32) is equal to the pentagon $C = \{0 \le R_1 \le 1,\ 0 \le R_2 \le 1,\ R_1 + R_2 \le 1.5\}$.

Example 2: Let $A_1 = A_2 = \{0,1,2\}$ and $B = \{1,2\}$

$$z_i = \begin{cases} x_i, & \text{if } x_i \neq 0 \text{ and } y_i = 0 \text{ and } y_{i-1} \neq 0 \\ y_i, & \text{if } y_i \neq 0 \text{ and } x_i = 0 \text{ and } x_{i-1} \neq 0 \\ (1/2,1/2), & \text{otherwise} \end{cases} \tag{33}$$

where $(1/2,1/2)$ indicates that z_i is equally likely to be 1 or 2.

In this channel it is necessary for the encoders to use some sort of time-sharing to achieve optimum rates because simultaneous zeros or nonzeros and consecutive zeros result in equally likely outputs. We take the following initial conditional distribution (this choice does not affect the capacity region but simplifies the proof):

$$z_1 = \begin{cases} x_1, & \text{if } y_1 = 0 \\ (1/2,1/2), & \text{if } y_1 \neq 0. \end{cases} \tag{34}$$

We will now investigate the maximum achievable rates when transmitter 1 (respectively, 2) sends nonzeros at odd-numbered (respectively, even-numbered) times (with no restrictions otherwise). Then, it follows from (33) and (34) that

$$z_{2k+1} = \begin{cases} x_{2k+1}, & \text{if } y_{2k+1} = 0 \\ (1/2,1/2), & \text{if } y_{2k+1} \neq 0 \end{cases} \tag{35a}$$

$$z_{2k} = \begin{cases} y_{2k}, & \text{if } x_{2k} = 0 \\ (1/2,1/2), & \text{if } x_{2k} \neq 0 \end{cases} \tag{35b}$$

which means that the channel is actually decoupled into two identical memoryless channels whose capacity region is obtained as follows.

If (X_{2k}, Y_{2k}, Z_{2k}) are connected by (35b), then their mutual informations are easily shown to be given (in bits) by

$$I(X_{2k}; Z_{2k}|Y_{2k}) = h_b(P[X_{2k} = 0]/2) - P[X_{2k} = 0]$$
$$I(Y_{2k}; Z_{2k}|X_{2k}) = h_b(P[Y_{2k} = 1])P[X_{2k} \neq 0]$$
$$I(X_{2k}, Y_{2k}; Z_{2k}) = h_b(1/2 + P[X_{2k} \neq 0](1/2 - P[Y_{2k} = 1])) - P[X_{2k} = 0]$$

where $h_b(x) = -x \log x - (1-x)\log(1-x)$. All these mutual informations are maximized simultaneously by $P[Y_{2k} = 1] = P[Y_{2k} = 2] = 1/2$, and so the capacity region of the channel in (35b) is

$$\bigcup_{0 \le p \le 1} \left\{ 0 \le R_1 \le h_b\left(\frac{p}{2}\right) - p = \bigcup_{0 \le p \le 1} \left\{ 0 \le R_1 \le h_b\left(\frac{p}{2}\right) - p \right.\right.$$
$$\left. 0 \le R_1 \le 1 - p \quad 0 \le R_2 \le 1 - h_b\left(\frac{p}{2}\right) \right\}$$
$$\left. R_1 + R_2 \le 1 - p \right\} \tag{36}$$

and the capacity region of (35a) is obtained by interchanging R_1 and R_2 in (36). The sum of these two regions divided by 2 (since each channel is only used half the time) is found in Fig. 1. Another example where the capacity region of the multiple-access channel is explicitly computed is the symbol-asynchronous energy-constrained Gaussian channel ([21] is devoted to the evaluation of the limit characterizing the capacity region).

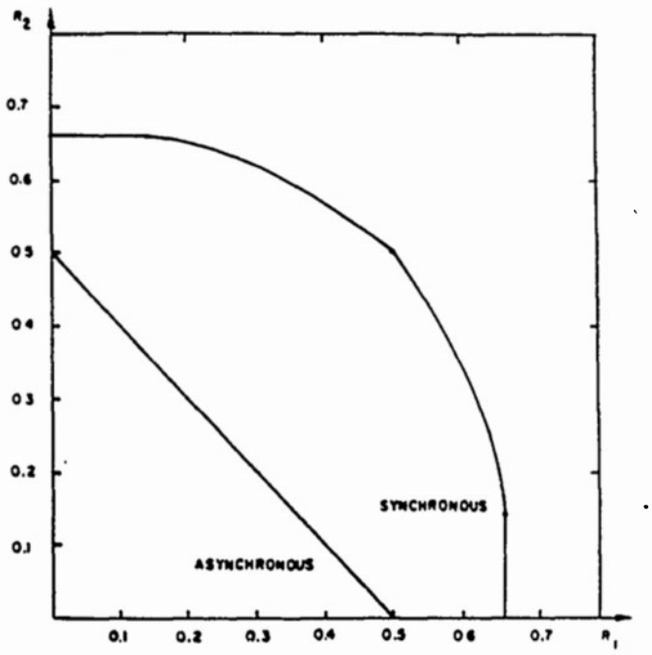

Fig. 1. Achievable regions with and without frame-synchronism of multiple-access channel in Example 2.

III. Frame-Asynchronous Capacity Region

Unlike frame-synchronous channels where it is enough to consider "one-shot" models in which each user transmits only one codeword, the (completely) frame-asynchronous multiple-access channel cannot be decoupled into independent blocks due to the overlap between consecutive codewords, and the optimum decoder needs to decode all messages simultaneously, i.e., all outputs are useful in making decisions about any particular codeword. Ideally, the goal would be to analyze a model with doubly infinite streams of codewords subject to an arbitrary shift. However, to formulate a well-posed problem, it is necessary (at least within the realm of channel block coding) to work with a finite number N, of transmitted codewords per user and then analyze the limiting behavior of the capacity as $N \to \infty$. Since the offset between both strings is arbitrary, the approach we take is to arrange the N codewords of each user in a ring (codeword N is followed by codeword 1) and to model the offset by an arbitrary relative rotation of both rings (Fig. 2). As $N \to \infty$, the radius of the ring becomes infinite, and the ring models the desired infinite codeword streams offset by an arbitrary shift because, for each output symbol, the boundary condition at infinity is irrelevant. As we will see and should expect (because of the finite memory the codeword boundaries become irrelevant

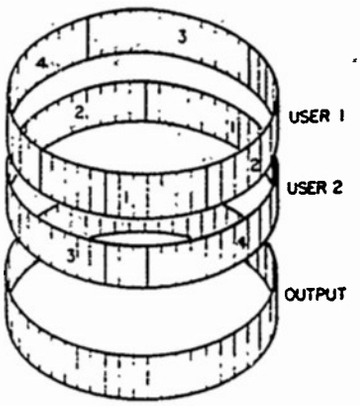

Fig. 2. Four-ring with codeword length equal to 12 and memory length $m = 2$.

as the codeword length goes to infinity), the capacity region (per channel use) of the N-ring does not depend on N, and therefore, it is not necessary to investigate its limiting behavior. The main disadvantage of an alternative linear arrangement of the N codewords is that, due to the lack of synchronism, not all the codewords overlap with the other user's stream, and those that do overlap have different decoding error probabilities depending on the offset and their relative location to the boundaries. Note that this problem can be partially avoided by restricting the shift not to exceed the length of a codeword, but since the total number of codewords is assumed finite, it can be argued that such an approach would assume a certain degree of cooperation between the transmitters.

Each transmitter encodes its N messages independently (each message is drawn independently from $\{1,\cdots,M_i\}$) and is not restricted to use the same code book for each message. While the receiver acquires the location of each codeword prior to decoding (this can be easily accomplished using synchronization prefixes), the messages are encoded without knowledge of the relative rotation. Therefore, the channel is a decoder-informed compound channel, which is equivalent, from the viewpoint of finding the capacity region, to a bank of parallel multiple-access channels (one per rotation value) sharing the same inputs.

Theorem 5: If (R_1, R_2) is an achievable rate pair for the N-ring, then

$$\begin{aligned}(R_1,R_2) &\in \text{closure}\left(\liminf_{n\to\infty}\frac{1}{nN}Q_{nN}^S\right)\\ &\subset \text{closure}\left(\limsup_{n\to\infty}\frac{1}{n}Q_n^S\right)\\ &= \text{closure}\left(\limsup_{n\to\infty}\frac{1}{n}C_n^S\right)\end{aligned}\tag{37}$$

where Q_n^S and C_n^S are defined as in (12) and (3), except that the union therein is taken only over n-dimensional distributions induced by stationary probability measures.

Proof: There are N (n, M_i) code books for the ith user and each of the N messages are encoded independently. Thus the (nN, M_i^N) juxtaposition code book of the ith user for the N-ring consists of the Cartesian product of the N (n, M_i) code books. If (R_1, R_2) is achievable, then for all $\epsilon > 0$, $\delta > 0$ and all n sufficiently large, there exists a $(nN, M_1^N, M_2^N, \epsilon)$ juxtaposition code for the N-ring such that

$$R_i \le \frac{\log M_i}{n} + \delta,$$

i.e., there exist N (n, M_1) code books for user 1 and N (n, M_2) codebooks for user 2 (which are independent of the offset) and a decoding strategy (which depends on the offset) such that the average (over the set of equiprobable messages) probability of error does not exceed ϵ regardless of the offset. Select one such code and denote the independent messages of both users by $(S_1,\cdots,S_N)$ and $(T_1,\cdots,T_N)$, respectively. Then, the Fano inequality implies that

$$H(S_1,\cdots,S_N|Z) \le \epsilon \log M_1^N + \log 2 \tag{38a}$$

$$H(T_1,\cdots,T_N|Z) \le \epsilon \log M_2^N + \log 2 \tag{38b}$$

$$H(S_1,\cdots,S_N,T_1,\cdots,T_N|Z) \le \epsilon \log M_1^N M_2^N + \log 2 \tag{38c}$$

where Z is the distribution of the totality of the outputs of the N-ring. If X and Y are the distributions of the A_1^{nN} and A_2^{nN} valued random variables resulting from the encoding of the messages by the selected code books, the lack of frame synchronism is modeled by assuming that the inputs to the N-ring are rotated versions of X and Y. If x is an M-vector, then $r_\tau(x)$ denotes an M-vector whose components coincide with those of x rotated by τ positions, where $\tau \in \{0,\cdots,M-1\}$, i.e.,

$$r_\tau(a_1,\cdots,a_M) = (a_{M-\tau+1},\cdots,a_M,a_1,\cdots,a_{M-\tau}).$$

Even though it is enough to consider a relative rotation of both rings, it is more convenient in the proof of the converse to allow a rotation of both input rings with respect to an arbitrary reference. Denoting the rotations by τ_1 and τ_2, the data-processing lemma implies that

$$\begin{aligned}I\left(r_{\tau_1}(X); Z, r_{\tau_2}(Y)\right) &\ge I(S_1,\cdots,S_N; Z)\\ &= H(S_1,\cdots,S_N) - H(S_1\cdots S_N|Z)\\ &\ge (1-\epsilon)\log M_1^N - \log 2\end{aligned}\tag{39}$$

where the last inequality follows from (38a) and with a slight abuse of notation we have denoted by $r_\tau(X)$ the probability measure that assigns the same mass to $r_\tau(a_1,\cdots,a_M)$ as X assigns to $(a_1,\cdots,a_M)$. From the independence of X and Y and (39), we have that

$$I\left(r_{\tau_1}(X); Z|r_{\tau_2}(Y)\right) \ge (1-\epsilon)\log M_1^N - \log 2.$$

However, since this is true regardless of the actual value of the offsets τ_1 and τ_2 we can write

$$\begin{aligned}(1-\epsilon)\log M_1^N - \log 2 &\le \frac{1}{nN}\sum_{\tau_2=0}^{nN-1} I\left(r_{\tau_1}(X); Z|r_{\tau_2}(Y)\right)\\ &= I\left(r_{\tau_1}(X); Z|c(Y)\right)\end{aligned}\tag{40}$$

where the probability measure $c(Y)$ is equal to the follow-

ing mixture of the probability measures $r_{\tau_2}(Y)$, $\tau_2 = 0,\cdots,nN-1$:

$$c(Y) = \frac{1}{nN}\sum_{\tau=0}^{nN-1} r_\tau(Y), \tag{41}$$

and the second equation in (40) follows from the fact that the distribution of the conditioning random variable enters linearly in the definition of conditional mutual information.

An M-dimensional probability measure p will be referred to as circulant if $r_\tau(p) = p$ for all τ, and we will say that p is stationary if, for any subset $\{i_1 \cdots i_t\} \subset \{1,\cdots,M\}$ and shift $s > 0$ such that $\{i_1+s,\cdots,i_t+s\} \subset \{1,\cdots,M\}$,

$$p(a_{i_1},\cdots,a_{i_t}) = p(a_{i_1+s},\cdots,a_{i_t+s})$$

For any probability measure p, $c(p)$ (defined in (41)) is a circulant probability measure. To see this note that, for any $\lambda \in \{0,\cdots,M\}$,[5]

$$r_\lambda(c(p)) = \frac{1}{M}\sum_{\tau=0}^{M-1} r_{(\tau+\lambda)_M}(p) = \frac{1}{M}\sum_{\tau=0}^{M-1} r_\tau(p) = c(p).$$

Furthermore, it is easy to check that an M-dimensional circulant probability measure is an M-dimensional stationary probability measure. Now, since (40) holds for all $\tau_1 \in \{0,\cdots,nN-1\}$,

$$(1-\epsilon)\log M_1^N - \log 2 \le \frac{1}{nN}\sum_{\tau_1=0}^{nN-1} I\big(r_{\tau_1}(X); Z|c(Y)\big) \le I(c(X); Z|c(Y)) \tag{42}$$

where the second inequality follows from the concavity of mutual information. Proceeding in a similar way we obtain from (38)

$$(1-\epsilon)\log M_2^N - \log 2 \le I(c(Y); Z|c(X)) \tag{43}$$

and

$$(1-\epsilon)\log M_1^N M_2^N - \log 2 \le I(c(X), c(Y); Z). \tag{44}$$

If $m-1$ consecutive components of Z are discarded, then we have a channel analogous to an nN-block channel whose B^{nN-m+1} valued output random variable is denoted by Z_m^{nN}. Then, the following upper bounds follow in a way similar to (18)

$$I(c(X); Z|c(Y)) \le I\big(c(X); Z_m^{nN}|c(Y)\big) + (m-1)\log|B| \tag{45a}$$

$$I(c(Y); Z|c(X)) \le I\big(c(Y); Z_m^{nN}|c(X)\big) + (m-1)\log|B| \tag{45b}$$

$$I(c(X), c(Y); Z) \le I\big(c(X), c(Y); Z_m^{nN}\big) + (m-1)\log|B|. \tag{45c}$$

Finally, it follows from (38), (42)–(45) and the stationarity of $c(X)$ and $c(Y)$ that

$$[1-\epsilon](R_1-\delta, R_2-\delta) - \left[\log\frac{2}{nN} + \frac{m-1}{nN}\log|B|\right](1,1) \in \frac{1}{nN}Q_{nN}^S.$$

Thus if $n > [\log 2 + (m-1)\log|B|]/[\delta(1-\epsilon)]$, then

$$[1-\epsilon](R_1-2\delta, R_2-2\delta) \in \frac{1}{nN}Q_{nN}^S$$

which implies that

$$[1-\epsilon](R_1-2\delta, R_2-2\delta) \in \liminf_{n\to\infty}\frac{1}{nN}Q_{nN}^S.$$

However, since ϵ and δ are arbitrarily small, (R_1, R_2) has to be a limit point of a sequence of points belonging to $\liminf_{n\to\infty}(1/nN)Q_{nN}^S$, and (37) follows.

Theorem 6: The following set is an achievable region for the N-ring:

$$\text{closure}\left(\bigcup_{\substack{\mu_X,\mu_Y \\ \text{stationary}}} \{(R_1,R_2): 0 \le R_1 \le I(\mu_X;\mu_Z|\mu_Y),\ 0 \le R_2 \le I(\mu_Y;\mu_Z|\mu_X),\ R_1+R_2 \le I(\mu_X,\mu_Y;\mu_Z)\}\right). \tag{46}$$

where

$$I(\mu_X;\mu_Z|\mu_Y) = \lim_{n\to\infty}\frac{1}{n}I(X^n; Z^n|Y^n) \tag{47a}$$

$$I(\mu_Y;\mu_Z|\mu_X) = \lim_{n\to\infty}\frac{1}{n}I(Y^n; Z^n|X^n) \tag{47b}$$

$$I(\mu_X,\mu_Y;\mu_Z) = \lim_{n\to\infty}\frac{1}{n}I(X^n, Y^n; Z^n) \tag{47c}$$

and X^n, Y^n are the n-dimensional distributions induced by the stationary probability measures μ_X, μ_Y, and Z^n is the output of the frame-synchronous multiple-access channel with memory when the inputs are independent with distributions X^n and Y^n.

Proof: The existence of the limits in (47) is an easy consequence of the stationarity of the inputs, the time-invariance of the channel, and the existence of entropy rate for any discrete stationary process (e.g., [9]). The symbols transmitted by each user will be denoted by

$$x = \{x_k(i),\ k=1,\cdots,N,\ i=1,\cdots,n\}$$
$$y = \{y_k(i),\ k=1,\cdots,N,\ i=1,\cdots,n\}$$

where n is the codeword length and N is the number of codewords in the ring. Similarly, the output symbols are labeled by

$$z = \{z_k(i),\ k=1,\cdots,N,\ i=1,\cdots,n\}.$$

If the users were frame-synchronous, then $z_k(i)$ would depend on $\{x_k(i-j)$ (or $x_{(k-1)_N}(i-j+n)$ if $i \le j)\}_{j=0}^{m-1}$ and $\{y_k(i-j)$ (or $y_{(k-1)_N}(i-j+n)$ if $i \le j)\}_{j=0}^{m-1}$. The lack of frame synchronism introduces a relative rotation of

[5] $(i)_M \in \{1,\cdots,M\}$ is equal to the remainder $i - qM$ where q is an integer.

the rings which can be quantified by $s \in \{0, \cdots, N-1\}$, the number of codewords shifted, and $r \in \{0, \cdots, n-1\}$, the rotation modulo the codeword length. More precisely, the input $x_k(i)$ is aligned with $\tilde{y}_k(i)$, defined by

$$\tilde{y}_r(j) = \begin{cases} y_{(l-s)_N}(j-r), & r < j \le n \\ y_{(l-s-1)_N}(j-r+n), & 1 \le j \le r \end{cases}$$

or, equivalently,

$$y_k(i) = \begin{cases} \tilde{y}_{(s+k)_N}(r+i), & i+r \le n \\ \tilde{y}_{(s+k+1)_N}(r+i-n), & i+r > n. \end{cases}$$

We will now fix an integer $l \ge 0$ independent of all other parameters and force the decoder to discard the following output values:

$$z_k(i),\ i \notin I;\ I = \{l+m, \cdots r\} \cup \{l+m+r, \cdots, n\}$$

which corresponds to discarding the $l+m-1$ symbols following the beginning of each received codeword. Note that if l is large enough, $I = \varnothing$; however, we will eventually be interested only in the asymptotic behavior as $n \to \infty$, in which case I is practically identical to $\{1, \cdots, n\}$. Note further that the relative shift r is allowed to be any integer $\{0, \cdots, n-1\}$, and so the cardinality of each of the two components of I may grow linearly in n. We may rearrange the codewords of Fig. 2 in the matrix form shown in Fig. 3. In this figure, each codeword of user 1 occupies a single row, whereas each codeword of user 2 occupies two consecutive rows. The blacked-out outputs correspond to the $l+m-1$ symbols following the beginning of each codeword which are discarded by the decoder. In connection with this figure, it is useful to introduce the following notation

$$\begin{aligned} x_k^L &= \{x_k(i),\ i = l+1, \cdots r\} \\ x_k^R &= \{x_k(i),\ i = l+r+1, \cdots, n\} \\ \tilde{y}_k^L &= \{\tilde{y}_k(i),\ i = l+1, \cdots, r\} \\ \tilde{y}_k^R &= \{\tilde{y}_k(i),\ i = l+r+1, \cdots, n\} \\ \tilde{z}_k^L &= \{\tilde{z}_k(i),\ i = l+m, \cdots, r\} \\ \tilde{z}_k^R &= \{\tilde{z}_k(i),\ i = l+r+m, \cdots, n\} \\ \tilde{z} &= \{z_k(i),\ k = 1, \cdots, N,\ i \in I\} \\ &= \{(z_k^L, z_k^R),\ k = 1, \cdots, N\}. \end{aligned}$$

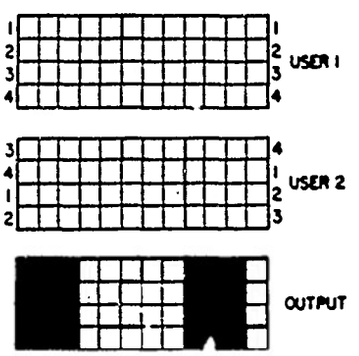

Fig. 3. Rearrangement of codewords in Fig. 2 showing discarded outputs ($l = 2$ and $m = 2$).

Then, the definition of the multiple-access channel with finite memory (10) implies that

$$\begin{aligned} p_{\tilde{Z}|X\tilde{Y}}(\tilde{z}|x, \tilde{y}) &= \prod_{k=1}^{N} \prod_{i \in I} p_c\big(z_k(i)|x_k(i-m+1), \cdots, \\ &\qquad x_k(i), \tilde{y}_k(i-m+1), \cdots, \tilde{y}_k(i)\big) \\ &= \prod_{k=1}^{N} p_{Z_k^L|X_k^L\tilde{Y}_k^L}\big(z_k^L|x_k^L, \tilde{y}_k^L\big) \\ &\qquad \cdot p_{Z_k^R|X_k^R\tilde{Y}_k^R}\big(z_k^R|x_k^R, \tilde{y}_k^R\big) \end{aligned} \tag{48}$$

which implies that the output subblocks $\{z_k^L, z_k^R,\ k = 1, \cdots, N\}$ are conditionally independent given the inputs and only depend on their corresponding input subblocks. Notice also that the inputs with indices $i \in \{1, \cdots, l\} \cup \{r+1, \cdots, r+l\}$ do not affect any outputs used by the decoder.

We now proceed to show that for any pair (μ_X, μ_Y) of stationary l-dependent measures defined on the infinite sequences drawn from A_1 and A_2, respectively, the following pentagon is achievable

$$\begin{aligned} C(\mu_X, \mu_Y) = \{(R_1, R_2):\ & 0 \le R_1 \le I(\mu_X; \mu_Z|\mu_Y) \\ & 0 \le R_2 \le I(\mu_Y; \mu_Z|\mu_Z) \\ & R_1 + R_2 \le I(\mu_X, \mu_Y; \mu_Z)\}. \end{aligned}$$

To this end, we must show that for any fixed $(R_1, R_2) \in C(\mu_X, \mu_Y)$, $\epsilon > 0$, and $\gamma > 0$, and all sufficiently large n there exist $(nN, M_1^N, M_2^N, \epsilon)$ juxtaposition codes for the N-ring such that

$$\frac{\log M_i}{n} \ge R_i - \gamma, \qquad i = 1, 2,$$

i.e., there exist N (n, M_i) code books for user i (independent of the relative rotation), and a decoding rule (possibly dependent of the relative rotation) such that the probability that any of the N messages transmitted by each user are decoded incorrectly is not higher than ϵ. The codes are chosen as follows.

Random Coding: The N code books of user i are denoted by $\{f_{ij}: \{1, \cdots, M_i\} \to A_i^n\}_{j=1}^N$ and are the outcomes of random selection where each codeword in each code book is independently selected with probability

$$\begin{aligned} p_{F_{1,j}(m)}(a_1, \cdots, a_n) &= p_{X^n}(a_1, \cdots, a_n) \\ &\qquad m \in \{1, \cdots, M_1\} \\ p_{F_{2,j}(m)}(b_1, \cdots, b_n) &= p_{Y^n}(b_1, \cdots, b_n) \\ &\qquad m \in \{1, \cdots, M_2\} \end{aligned}$$

where X^n and Y^n are the independent n-dimensional distributions induced by (μ_X, μ_Y). The overall code book

of user i resulting from the juxtaposition of the foregoing N code books is denoted by f_i: $\{1,\cdots,M_i\}^N \to A_i^{nN}$.

Decoding: The decoder performs simultaneous decoding of the N messages transmitted by both users, upon observing $\tilde{z}$ and the rotation (r,s), in the knowledge of the code books f_1 and f_2. The decoder selects the messages $(m_1,m_2) \in \{1,\cdots,M_1\}^N \times \{1,\cdots,M_2\}^N$ if (m_1,m_2) is the unique pair that satisfies

$$(f_1(m_1), f_2(m_2), \tilde{z}) \in J(n,\delta) \tag{49}$$

where $0<\delta<\gamma$, and it outputs a decoding error if there are zero or more than one such pairs. The set $J(n,\delta)$ in (49) is the set of jointly typical sequences according to three criteria:

$$J(n,\delta) = J_1(n,\delta) \cap J_2(n,\delta) \cap J_3(n,\delta)$$

$$J_1(n,\delta) = \left\{(x,\tilde{y},\tilde{z}): \left|\frac{1}{n} i_1(x,\tilde{y},\tilde{z}) - \frac{1}{n} I(X;\tilde{Z}|\tilde{Y})\right| \le \delta\right\} \tag{50a}$$

$$J_2(n,\delta) = \left\{(x,\tilde{y},\tilde{z}): \left|\frac{1}{n} i_2(x,\tilde{y},\tilde{z}) - \frac{1}{n} I(\tilde{Y};\tilde{Z}|X)\right| \le \delta\right\} \tag{50b}$$

$$J_3(n,\delta) = \left\{(x,\tilde{y},\tilde{z}): \left|\frac{1}{n} i_3(x,\tilde{y},\tilde{z}) - \frac{1}{n} I(X,\tilde{Y};\tilde{Z})\right| \le \delta\right\} \tag{50c}$$

where

$$i_1(x,\tilde{y},\tilde{z}) = \log \frac{p_{\tilde{Z}|X\tilde{Y}}(\tilde{z}|x,\tilde{y})}{p_{\tilde{Z}|\tilde{Y}}(\tilde{z}|\tilde{y})} \tag{51a}$$

$$i_2(x,\tilde{y},\tilde{z}) = \log \frac{p_{\tilde{Z}|X\tilde{Y}}(\tilde{z}|x,\tilde{y})}{p_{\tilde{Z}|X}(\tilde{z}|x)} \tag{51b}$$

$$i_3(x,\tilde{y},\tilde{z}) = \log \frac{p_{\tilde{Z}|X\tilde{Y}}(\tilde{z}|x,\tilde{y})}{p_{\tilde{Z}}(\tilde{z})}. \tag{51c}$$

Note that the expected values of the functions in (51) evaluated with the distribution $(X,\tilde{Y},\tilde{Z})$ are equal to the mutual informations appearing in (50). We can decompose the functions in (51) taking advantage of the assumed l-dependence of the inputs, which implies that the random variables $\{X_k^L, X_k^R,\ k=1,\cdots,N\}$ (and $\{\tilde{Y}_k^L, \tilde{Y}_k^R,\ k=1,\cdots,N\}$) are independent. Then we obtain via (48) that

$$i_1(x,\tilde{y},\tilde{z}) = \sum_{k=1}^{N} \log \frac{p_{\tilde{Z}_k^L|X_k^L\tilde{Y}_k^L}(\tilde{z}_k^L|x_k^L,\tilde{y}_k^L)}{p_{\tilde{Z}_k^L|\tilde{Y}_k^L}(\tilde{z}_k^L|\tilde{y}_k^L)} + \frac{p_{\tilde{Z}_k^R|X_k^R\tilde{Y}_k^R}(\tilde{z}_k^R|x_k^R,\tilde{y}_k^R)}{p_{\tilde{Z}_k^R|\tilde{Y}_k^R}(\tilde{z}_k^R|\tilde{y}_k^R)} \tag{52a}$$

$$i_2(x,\tilde{y},\tilde{z}) = \sum_{k=1}^{N} \log \frac{p_{\tilde{Z}_k^L|X_k^L\tilde{Y}_k^L}(\tilde{z}_k^L|x_k^L,\tilde{y}_k^L)}{p_{\tilde{Z}_k^L|X_k^L}(\tilde{z}_k^L|x_k^L)} + \frac{p_{\tilde{Z}_k^R|X_k^R\tilde{Y}_k^R}(\tilde{z}_k^R|x_k^R,\tilde{y}_k^R)}{p_{\tilde{Z}_k^R|X_k^R}(\tilde{z}_k^R|x_k^R)} \tag{52b}$$

$$i_3(x,\tilde{y},\tilde{z}) = \sum_{k=1}^{N} \log \frac{p_{\tilde{Z}_k^L|X_k^L\tilde{Y}_k^L}(\tilde{z}_k^L|x_k^L,\tilde{y}_k^L)}{p_{\tilde{Z}_k^L}(\tilde{z}_k^L)} + \frac{p_{\tilde{Z}_k^R|X_k^R\tilde{Y}_k^R}(\tilde{z}_k^R|x_k^R,\tilde{y}_k^R)}{p_{\tilde{Z}_k^R}(\tilde{z}_k^R)}. \tag{52c}$$

Taking expected values of (52) with respect to $(X,\tilde{Y},\tilde{Z})$ and recalling that the inputs are stationary we obtain that

$$I(X;\tilde{Z}|\tilde{Y}) = NI(X^n;\tilde{Z}^n|\tilde{Y}^n) \tag{53a}$$

$$I(\tilde{Y};\tilde{Z}|X) = NI(\tilde{Y}^n;\tilde{Z}^n|X^n) \tag{53b}$$

$$I(X,\tilde{Y};\tilde{Z}) = NI(X^n,\tilde{Y}^n;\tilde{Z}^n) \tag{53c}$$

where $\tilde{Z}^n$ and $\tilde{Y}^n$ denote $\{Z_1(i),\ i\in I\}$ and $\{\tilde{Y}_1(i),\ i=1,\cdots,n\}$, respectively. Furthermore, since $\{Z_1(i),\ i\in I\}$ depends on the inputs only through X_1^L, X_1^R, $\tilde{Y}_1^L$, and $\tilde{Y}_1^R$, and since μ_γ is stationary and l-dependent, $\{\tilde{Y}_1^L, \tilde{Y}_1^R\}$ has the same distribution as $\{Y_1^L, Y_1^R\}$ and therefore (53) can be written as

$$I(X;\tilde{Z}|\tilde{Y}) = NI(X^n;\tilde{Z}^n|Y^n) \tag{54a}$$

$$I(\tilde{Y};\tilde{Z}|X) = NI(Y^n;\tilde{Z}^n|X^n) \tag{54b}$$

$$I(X,\tilde{Y};\tilde{Z}) = NI(X^n,Y^n;\tilde{Z}^n). \tag{54c}$$

The probability that the transmitted messages $(S_1,S_2) = (m_1,m_2)$ are not decoded correctly given that (f_1,f_2) are the chosen code books is

$$e_{m_1m_2}(f_1,f_2) = P\Big[\big(F_1(m_1), F_2(m_2), \tilde{Z}\big) \notin J(n,\delta) \text{ or } \exists (m_1',m_2') \ne (m_1,m_2) \text{ such that } \big(F_1(m_1'), F_2(m_2'), \tilde{Z}\big) \in J(n,\delta) \big| (F_1,F_2) = (f_1,f_2), (S_1,S_2) = (m_1,m_2)\Big].$$

Averaging over the random selection of code books and invoking the union bound, we obtain

$$\begin{aligned} E\left[e_{m_1m_2}(F_1,F_2)\right] \le\ & P\left[\big(F_1(m_1),F_2(m_2),\tilde{Z}\big) \notin J(n,\delta) \big| (S_1,S_2) = (m_1,m_2)\right] \\ & + \sum_{m_1' \ne m_1} \sum_{m_2' \ne m_2} P\left[\big(F_1(m_1'),F_2(m_2'),\tilde{Z}\big) \in J_3(n,\delta) \big| (S_1,S_2) = (m_1,m_2)\right] \\ & + \sum_{m_1' \ne m_1} P\left[\big(F_1(m_1'),F_2(m_2),\tilde{Z}\big) \in J_1(n,\delta) \big| (S_1,S_2) = (m_1,m_2)\right] \\ & + \sum_{m_2' \ne m_2} P\left[\big(F_1(m_1),F_2(m_2'),\tilde{Z}\big) \in J_2(n,\delta) \big| (S_1,S_2) = (m_1,m_2)\right]. \end{aligned} \tag{55}$$

The first term in the right side of (55) is smaller than $\epsilon/2$ for sufficiently large n because

$$\lim_{n\to\infty} P\left[\left(F_1(m_1), F_2(m_2), \tilde{Z}\right) \in J_k(n,\delta) \mid (S_1, S_2) = (m_1, m_2)\right] = 1, \qquad k = 1,2,3. \quad (56)$$

This holds because the inputs and output are jointly stationary and ergodic [4] (the output is $(l+m+1)$-dependent) and therefore the Shannon–McMillan theorem (see e.g., [9]) implies that each of the $2N$ terms in the right sides of (52) converges in probability (when scaled by n) to its expected value (which may be zero if, for example, r remains finite as $n\to\infty$). (Notice that this holds even though each output subblock z_k^L (or z_k^R) has $m-1$ fewer elements than the corresponding input subblocks x_k^L and $\tilde{y}_k^L$ (or x_k^R and $\tilde{y}_k^R$), since convergence is not affected by any fixed number of elements.)

To investigate the behavior of the second term on the right side of (55), we will introduce the independent random vectors U and V defined on A_1^{nN} and A_2^{nN}, respectively, whose distributions are p_X and $p_{\tilde{Y}}$, but which, unlike X and $\tilde{Y}$, are independent of $\tilde{Z}$, i.e.,

$$p_{UV\tilde{Z}}(x,\tilde{y},\tilde{z}) = p_X(x)p_{\tilde{Y}}(\tilde{y})p_{\tilde{Z}}(\tilde{z}) = p_{X\tilde{Y}\tilde{Z}}(x,\tilde{y},\tilde{z})\exp\left(-i_3(x,\tilde{y},\tilde{z})\right). \quad (57)$$

If $(x,\tilde{y},\tilde{z}) \in J_3(n,\delta)$, however, then (50c) implies that

$$\exp\left(-i_3(x,\tilde{y},\tilde{z})\right) \le \exp\left(-I(X,\tilde{Y};\tilde{Z}) + n\delta\right)$$

and, consequently,

$$P\left[\left(F_1(m_1'), F_2(m_2'), \tilde{Z}\right) \in J_3(n,\delta) \mid (S_1,S_2) = (m_1,m_2)\right] = \sum_{(x,\tilde{y},\tilde{z}) \in J_3(n,\delta)} p_{UV\tilde{Z}}(x,\tilde{y},\tilde{z}) \le \exp\left(-I(X,\tilde{Y};\tilde{Z}) + n\delta\right). \quad (58)$$

Proceeding similarly with the third and fourth terms on the right side of (55) we obtain that, for sufficiently large n,

$$E\left[e_{m_1m_2}(F_1,F_2)\right] \le M_1^N \exp\left(-I(X;\tilde{Z}|\tilde{Y}) + n\delta\right) + M_2^N \exp\left(-I(\tilde{Y};\tilde{Z}|X) + n\delta\right) + M_1^N M_2^N \exp\left(-I(X,\tilde{Y};\tilde{Z}) + n\delta\right) + \frac{\epsilon}{2}. \quad (59)$$

Thus if M_1 and M_2 grow sufficiently slowly with n we will be able to show that for large n the right side of (59) does not exceed ϵ. Specifically, we choose M_1 and M_2 to satisfy

$$R_i - \frac{\gamma}{N} < \frac{\log M_i}{n} \le R_i - \frac{\delta+\gamma}{2N}, \qquad i = 1,2. \quad (60)$$

(This choice is possible for all sufficiently large n.) Then, (59) is further upper-bounded by

$$E\left[e_{m_1m_2}(F_1,F_2)\right] \le \exp\left(nN\left[R_1 - \frac{1}{nN}I(X;\tilde{Z}|\tilde{Y})\right] - n\frac{\gamma-\delta}{2}\right) + \exp\left(nN\left[R_2 - \frac{1}{nN}I(\tilde{Y};\tilde{Z}|X)\right] - n\frac{\gamma-\delta}{2}\right) + \exp\left(nN\left[R_1+R_2 - \frac{1}{nN}I(X,\tilde{Y};\tilde{Z})\right] - n\gamma\right) + \frac{\epsilon}{2}. \quad (61)$$

Using (54) and recalling that $\delta < \gamma$, it is seen that if

$$0 \le R_1 \le \liminf_{n\to\infty} \frac{1}{n} I(X^n;\tilde{Z}^n|Y^n) \quad (62a)$$

$$0 \le R_2 \le \liminf_{n\to\infty} \frac{1}{n} I(Y^n;\tilde{Z}^n|X^n) \quad (62b)$$

$$R_1 + R_2 \le \liminf_{n\to\infty} \frac{1}{n} I(X^n,Y^n;\tilde{Z}^n), \quad (62c)$$

then the right side of (61) does not exceed ϵ for sufficiently large n. Therefore, at least one realization of the code books must exist that results in probability of error better than ϵ, and so the pentagon in (62) is achievable. Actually, that region coincides with $C(\mu_X,\mu_Y)$ because $\tilde{Z}^n$ was obtained by discarding $2(l+m-1)$ elements from $Z^n = \{Z_1(i), i = 1,\cdots,n\}$ (which is the output of the frame-synchronous channel with inputs X^n and Y^n) and therefore (cf. (18))

$$I(X^n;\tilde{Z}^n|Y^n) \le I(X^n;Z^n|Y^n) \le I(X^n;\tilde{Z}^n|Y^n) + 2(l+m-1)\log|B| \quad (63a)$$

$$I(Y^n;\tilde{Z}^n|X^n) \le I(Y^n;Z^n|X^n) \le I(Y^n;\tilde{Z}^n|X^n) + 2(l+m-1)\log|B| \quad (63b)$$

$$I(X^n,Y^n;\tilde{Z}^n) \le I(X^n,Y^n;Z^n) \le I(X^n,Y^n;\tilde{Z}^n) + 2(l+m-1)\log|B|. \quad (63c)$$

Hence we may replace $\tilde{Z}^n$ by Z^n in (62), obtaining the limits of (47). Thus we have shown that the region

$$\text{closure}\left(\bigcup_{l\ge 0} \bigcup_{\substack{\mu_X,\mu_Y \\ \text{stationary } l\text{-dep}}} \{(R_1,R_2): 0 \le R_1 \le I(\mu_X;\mu_Z|\mu_Y),\ 0 \le R_2 \le I(\mu_Y;\mu_Z|\mu_X),\ R_1+R_2 \le I(\mu_X,\mu_Y;\mu_Z)\}\right) \quad (64)$$

is achievable.

It remains to show that the restriction to l-dependent input distributions can be dropped without changing the region in (64). We will do so in three steps where we show that the union can be written over 1) B-processes, 2) ergodic stationary processes, and finally, 3) stationary processes.

Step 1: The B-processes are an important class of stationary ergodic discrete-time random processes (introduced by Ornstein [16]) that can be defined as the outputs of the time-invariant systems driven by independent identically distributed (i.i.d.) inputs. This is, in effect, a mixing condition requiring that the influence of the sufficiently distant past becomes negligible. It was shown by Ornstein [16] that the closure in the $\bar{d}$-metric of the stationary l-dependent processes is equal to the set of B-processes. The $\bar{d}$-metric between two stationary ergodic measures μ and $\bar{\mu}$ is equal to the minimum percentage of time samples we need to change a representative realization of μ to make it look like a representative realization of $\bar{\mu}$. Due to the finite memory of the channel, it is easy to show that if a sequence of stationary ergodic input measures $\mu_X^{(k)}, \mu_Y^{(k)}$ converges in the $\bar{d}$-metric to μ_X and μ_Y, then the corresponding output measures also converge in the $\bar{d}$-metric, because one way to generate a representative sequence of μ_Z is by modifying representative strings of $\mu_X^{(k)}$ and $\mu_Y^{(k)}$ to get representative strings of μ_X and μ_Y without changing the output samples unaffected by those modifications. Therefore, $\bar{d}(\mu_Z^{(k)}, \mu_Z) \le m[\bar{d}(\mu_X^{(k)}, \mu_X) + \bar{d}(\mu_Y^{(k)}, \mu_Y)]$ since each input value affects at most m output values. Now, since the entropy rate is a continuous function of the stationary measure under the $\bar{d}$-metric [19] and the three constraints in (64) can be written as

$$I(\mu_X; \mu_Z|\mu_Y) = H(\mu_Y, \mu_Z) + H(\mu_X) - H(\mu_X, \mu_Y, \mu_Z) \tag{65a}$$

$$I(\mu_Y; \mu_Z|\mu_X) = H(\mu_X, \mu_Z) + H(\mu_Y) - H(\mu_X, \mu_Y, \mu_Z) \tag{65b}$$

$$I(\mu_X, \mu_Y; \mu_Z) = H(\mu_X) + H(\mu_Y) + H(\mu_Z) - H(\mu_X, \mu_Y, \mu_Z), \tag{65c}$$

the region (64) is unchanged if we enlarge the set of stationary l-dependent processes to its closure, the set of B-processes.

Step 2: The stationary mixing multistep Markov processes are B-processes (the mixing condition essentially rules out processes with periodicities), whose closure in the weak topology is the set of stationary ergodic processes [10, p. 360], where we say that $\mu^{(k)}$ converges weakly to μ if, for all $n > 0$, the n-dimensional distribution induced by $\mu^{(k)}$ converges to that of μ. Again, due to the finite memory of the channel it is easy to show that if $\mu_X^{(k)}, \mu_Y^{(k)}$ converge weakly to μ_X and μ_Y, then the corresponding output measures also converge weakly, because the n-dimensional distribution induced by $\mu_Z^{(k)}$ depends linearly on the $n+m-1$-dimensional distributions of $\mu_X^{(k)}$ and $\mu_Y^{(k)}$. Furthermore, since the entropy rate is a lower semicontinuous function in the weak topology, we can write

$$\liminf_{k\to\infty} H(\mu_Z^{(k)}) \ge H(\mu_Z)$$

and (cf. [17])

$$\liminf_{k\to\infty} H(\mu_Z^{(k)}|\mu_Y^{(k)}) \ge H(\mu_Z|\mu_Y),$$

and

$$\lim_{k\to\infty} H(\mu_Z^{(k)}|\mu_X^{(k)}, \mu_Y^{(k)}) = H(\mu_Z|\mu_X, \mu_Y),$$

since the latter expression is linear in the conditioning measures. This implies that the union appearing in the achievable region can indeed be extended to the stationary ergodic measures.

Step 3: This step has a well-known counterpart in the solution of the capacity of single-user channels with memory (cf. [11, sec. III]). There, the ergodic assumption is needed to invoke the Shannon–McMillan theorem in the proof of the direct theorem, whereas the usual converse techniques upper-bound capacity by the minimum of mutual information rates over all stationary inputs. A proof that the lower and upper bounds thus obtained coincide was given by Parthasarathy [17] using the ergodic decomposition theorem. Even though in the multiuser case capacity is not given as the maximization of a scalar function, we can use Parthasarathy's result by noticing that all we need to show is that for every $0 \le \alpha \le 1$ (cf. [21])

$$\sup_{\substack{\mu_X\mu_Y \\ \text{stationary} \\ \text{ergodic}}} G_\alpha(\mu_X, \mu_Y) = \sup_{\substack{\mu_X\mu_Y \\ \text{stationary}}} G_\alpha(\mu_X, \mu_Y) \tag{66}$$

where

$$G_\alpha(\mu_X, \mu_Y) = \max_{\substack{0 \le R_1 \le I(\mu_X; \mu_Z|\mu_Y) \\ 0 \le R_2 \le I(\mu_Y; \mu_Z|\mu_X) \\ R_1 + R_2 \le I(\mu_X, \mu_Y; \mu_Z)}} \alpha R_1 + (1-\alpha) R_2$$

$$= \begin{cases} (2\alpha - 1) I(\mu_X; \mu_Z|\mu_Y) + (1-\alpha) I(\mu_X, \mu_Y; \mu_Z), & 1/2 \le \alpha \le 1 \\ (1-2\alpha) I(\mu_Y; \mu_Z|\mu_X) + \alpha I(\mu_X, \mu_Y; \mu_Z), & 0 \le \alpha \le 1/2 \end{cases} \tag{67}$$

where the second equality holds because the maximization on the left side is attained at one of the two (Pareto) optimum vertices of the feasible pentagon (note that $I(\mu_X, \mu_Y; \mu_Z) \le I(\mu_X; \mu_Z|\mu_Y) + I(\mu_Y; \mu_Z|\mu_X)$). We may now fix $\alpha \in [0, 1/2]$, the other case being entirely parallel. We then obtain that for all stationary pairs μ_X, μ_Y

$$\begin{aligned} G_\alpha(\mu_X, \mu_Y) &= (1-\alpha) I(\mu_X, \mu_Y; \mu_Z) - (1-2\alpha) I(\mu_X; \mu_Z) \\ &= (1-\alpha) \int I(\mu_{X_x}, \mu_{Y_y}; \mu_Z)\, dP_X(x)\, dP_Y(y) \\ &\quad - (1-2\alpha) \int I(\mu_{X_x}; \mu_Z)\, dP_X(x) \\ &= \int G_\alpha(\mu_{X_x}, \mu_{Y_y})\, dP_X(x)\, dP_Y(y) \end{aligned} \tag{68}$$

where the second equation follows from Parthasarathy's representation theorem [17], and $\{\mu_{X_x}, x \in \Lambda_1\}$ and $\{\mu_{Y_y}, y \in \Lambda_2\}$ are the stationary ergodic measures in the ergodic decompositions of μ_X and μ_Y:

$$\mu_X(E) = \int_{\Lambda_1} \mu_{X_x}(E)\, dP_X(x) \tag{69a}$$

$$\mu_Y(F) = \int_{\Lambda_2} \mu_{Y_y}(F)\, dP_Y(y) \tag{69b}$$

for all measurable sets E and F. Notice that the only restriction of Parthasarathy's result is that the channel connecting input and output be stationary, and this is the case for the channel that connects (μ_X, μ_Y) with μ_Z, as well as the channel seen by the first user, which connects μ_X and μ_Z because both μ_Y and the multiple-access channel are stationary.

Finally, for a stationary pair μ_X, μ_Y to achieve a value of $G_a(\mu_X, \mu_Y)$ close to the supremum, there must exist $(x, y) \in \Lambda_1 \times \Lambda_2$ such that $G_a(\mu_{X_x}, \mu_{Y_y})$ is close to the supremum, because the average with respect to $P_X \times P_Y$ in (68) cannot be larger than each of its sample values. This fact shows (66) and completes the proof of the theorem.

We will now use Theorems 5 and 6 to find the frame-asynchronous capacity region of the examples we studied at the end of Section II.

Example 1: Since the frame-synchronous capacity region is achieved by stationary (i.i.d.) inputs, it remains the same if the users are frame-asynchronous. The same is true for the symbol-asynchronous Gaussian multiple-access channel [21] where the capacity region is achieved by stationary colored Gaussian processes. (In that case, the capacity region does depend, in general, on whether the transmitters are symbol-synchronous.)

Example 2: We will show first that the triangle $\{0 \le R_1, 0 \le R_2, R_1 + R_2 \le 0.5$ bit$\}$ (Fig. 1) is an outer bound to the frame-synchronous capacity region. In the second part of the proof we will show that it is achievable. From the definition of this channel (33), it is easy to compute the conditional entropy (in bits):

$$H(Z_i | X_{i-1}, X_i, Y_{i-1}, Y_i) = 1 - \beta_1\gamma_2 - \beta_2\gamma_1 \tag{70}$$

with

$$\gamma_1 = P[X(i) \neq 0] \qquad \beta_1 = P[X(i) = 0,\ X(i-1) \neq 0]$$
$$\gamma_2 = P[Y(i) \neq 0] \qquad \beta_2 = P[Y(i) = 0,\ Y(i-1) \neq 0]$$

where the foregoing probabilities are independent of the time i because X^n and Y^n are stationary. Since the outputs are independent conditioned on the inputs, we have

$$\begin{aligned} I(X^n, Y^n; Z_m^n) &\le \sum_{i=m}^{n} I(X_{i-1}, X_i, Y_{i-1}, Y_i; Z_i) \\ &\le \sum_{i=m}^{n} [1 - H(Z_i | X_{i-1}, X_i, Y_{i-1}, Y_i)] \\ &= (n-m+1)[\gamma_1\beta_2 + \gamma_2\beta_1], \end{aligned} \tag{71}$$

and since $\gamma_k + \beta_k \in [0,1]$ and $\gamma_k - \beta_k \in [0,1]$, we have $\beta_k \in [0, 1/2]$ and

$$\max_{\gamma_1\gamma_2\beta_1\beta_2} \gamma_1\beta_2 + \gamma_2\beta_1 = 1/2$$

which, together with (71), implies that the total capacity of the frame-asynchronous channel is bounded from above by

$$\lim_{n\to\infty} \frac{1}{n} \max_{\substack{X^n Y^n \\ \text{stationary}}} I(X^n, Y^n; Z_m^n) \le 1/2,$$

in contrast to the total capacity of the frame-synchronous channel, which is equal to 1 bit.

To show achievability of the triangle, we choose both input processes to be stationary and Markov with

$$P[X_i \neq 0 | X_{i-1} = 0] = 1$$
$$P[Y_i \neq 0 | Y_{i-1} = 0] = 1$$

(i.e., $\gamma_k = 1 - \beta_k$, $k = 1,2$). Then, as there are no consecutive input zeros, the channel is equivalent to the memoryless channel

$$z_i = \begin{cases} x_i, & \text{if } x_i \neq 0 \text{ and } y_i = 0 \\ y_i, & \text{if } y_i \neq 0 \text{ and } x_i = 0 \\ (1/2, 1/2), & \text{otherwise.} \end{cases} \tag{72}$$

Furthermore, we will only consider inputs whose nonzero values are independent and equally likely to be 1 or 2. Then the outputs are independent both unconditionally and conditioned on either input sequence:

$$H(Z_m^n) = n - m + 1 \tag{73}$$

$$\begin{aligned} H(Z_m^n | Y^n) &= \sum_{i=m}^{n} H(Z_i | Y^n) \\ &= \sum_{i=m}^{n} H(Z_i | Y_i) \\ &= (n-m+1)\left[1 - \gamma_2 + \gamma_2 h_b\left(\frac{\gamma_1}{2}\right)\right] \end{aligned} \tag{74}$$

where the second equation follows from (72) and the independence of X^n and Y^n. Then (70), (73), and (74) imply that

$$I(X^n; Z_m^n | Y^n) = (n-m+1) \cdot \left[\gamma_1(1-\gamma_2) + \gamma_2\left(h_b\left(\frac{\gamma_1}{2}\right) - \gamma_1\right)\right]$$

$$I(Y^n; Z_m^n | X^n) = (n-m+1) \cdot \left[\gamma_2(1-\gamma_1) + \gamma_1\left(h_b\left(\frac{\gamma_2}{2}\right) - \gamma_2\right)\right]$$

$$I(X^n, Y^n; Z_m^n) = (n-m+1) \cdot [\gamma_1(1-\gamma_2) + \gamma_2(1-\gamma_1)].$$

However, $h_b(\gamma_k/2) \ge \gamma_k$, and thus the following region is achievable:

$$\bigcup_{\substack{\gamma_i \in [1/2,1] \\ i=1,2}} \{(R_1, R_2): 0 \le R_1 \le \gamma_1(1-\gamma_2),\ 0 \le R_2 \le \gamma_2(1-\gamma_1)\} \tag{75}$$

ynich can be shown to coincide with the triangle $\{(R_1, R_2)$: $) \le R_1, 0 \le R_2, R_1 + R_2 \le 1/2\}$.

t can be seen that in this example full frame-synchro- ious capacity would be achieved if the encoders were nformed of the relative shift modulo 2, and without this nformation, they cannot do better than the frame- synchronous region. This points out that, in contrast to he memoryless channel, even a mild form of asynchro- nism where the shift may be only 0 or 1 reduces the apacity region. The reason is that in the memoryless nannel, a large time-scale type of cooperation (time-shar- ng) is enough to achieve capacity, whereas in a channel vith memory, the encoders may need to cooperate in a mall time-scale. Mild frame-asynchronism only precludes ooperation in the small time-scale.

REFERENCES

[1] R. Ahlswede, "Multi-way communication channels," in *Proc. 2nd nt. Symp. Information Theory*, Tsahkadsor, USSR, 1971, pp. 103–135.

[2] R. Ahlswede and P. Gacs, "Two contributions to information theory," *Coll. Math. Soc. J. Bolyai*, vol. 16, pp. 17–40, 1975.

[3] M. Bierbaum and H. M. Wallmeier, "A note on the capacity region of the multi-access channel," *IEEE Trans. Inform. Theory*, vol. IT-25, p. 484, July 1979.

[4] P. Billingsley, *Ergodic Theory and Information*. New York: Wiley, 1965.

[5] T. Cover, R. McEliece, and E. Posner, "Asynchronous multiple-access channel capacity," *IEEE Trans. Inform. Theory*, vol. IT-27, pp. 409–413, July 1981.

[6] I. Csiszar and J. Korner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. New York: Academic, 1981.

[7] R. L. Dobrushin, "Shannon's theorems for channels with synchronization errors," *Prob. Inform. Transmission*, vol. 3, pp. 11–26, 1967.

[8] A. Feinstein, "On the coding theorem and its converse for finite-memory channels," *Inform. Contr.*, vol. 2, pp. 25–44, 1959.

[9] R. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968.

[10] R. M. Gray, "Sliding-block source coding," *IEEE Trans. Inform. Theory*, vol. IT-21, pp. 357–368, July 1975.

[11] R. M. Gray and L. D. Davisson, "The ergodic decomposition of stationary discrete random processes," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 625–636, Sept. 1974.

[12] J. Y. N. Hui and P. A. Humblet, "The capacity region of the totally asynchronous multiple-access channels," *IEEE Trans. Inform. Theory*, vol. IT-31, pp. 207–216, Mar. 1985.

[13] J. L. Massey and P. Mathys, "The collision channel without feedback," *IEEE Trans. Inform. Theory*, vol. IT-31, pp. 192–204, Mar. 1985.

[14] E. van der Meulen, "A survey of multi-way channels in information theory, 1961–1976," *IEEE Trans. Inform. Theory*, vol. IT-23, Jan. 1977.

[15] P. Narayan and D. Snyder, "The two-user cutoff rate for an asynchronous and a synchronous multiple access channel are the ame," *IEEE Trans. Inform. Theory*, vol. IT-27, pp. 414–419, July 981.

[16] D. S. Ornstein, "An application of ergodic theory to probability heory," *Ann. Probab.*, vol. 1, pp. 43–58, 1973.

[17] K. R. Parthasarathy, "On the integral representation of the rate of transmission of a stationary channel," *Illinois J. Math.*, vol. 2, pp. 299–305, 1961.

[18] G. S. Poltyrev, "Coding in an asynchronous multiple-access channel," *Prob. Inform. Transmission*, pp. 12–21, July–Sept. 1983.

[19] P. Shields, *Theory of Bernoulli Shifts*. Chicago, IL: Univ. of Chicago Press, 1973.

[20] I. P. Tsaregradsky, "On the capacity of a stationary channel with finite memory," *Teor. Verojatnost. Preimenen*, vol. 3, pp. 84–96, 1958.

[21] S. Verdu, "The capacity region of the symbol-asynchronous Gaussian multiple-access channel," *IEEE Trans. Inform. Theory*, vol. 35, July 1989, to be published.

[22] J. Wolfowitz, *Coding Theorems of Information Theory*, 3rd ed. New York: Springer, 1978.

[23] R. Ahlswede, "Eight problems in information theory," in *Open Problems in Communication and Computation*, T. M. Cover and B. Gopinath, Eds. New York: Springer, 1987.

The Capacity Region of the Symbol-Asynchronous Gaussian Multiple-Access Channel

Sergio Verdú

Reprinted from
IEEE TRANSACTIONS ON INFORMATION THEORY
Vol. 35, No. 4, July 1989

The Capacity Region of the Symbol-Asynchronous Gaussian Multiple-Access Channel

SERGIO VERDÚ, SENIOR MEMBER, IEEE

Abstract—In the information theory of the multiple-access channel, two types of synchronism are usually assumed among the transmitters, namely, frame and symbol synchronism. Frame synchronism refers to the ability of the users to start the transmission of their codewords in unison. The issue of symbol synchronism arises in continuous-time channels in which each codeword symbol modulates a fixed assigned waveform; the channel is symbol synchronous if the users cooperate so that their symbol epochs coincide at the receiver. In practice symbol synchronism is harder to achieve, yet the only reported progress so far has been in the removal of the assumption of frame synchronism. It is shown that if the transmitters are assigned the same waveform, symbol asynchronism has no effect on the two-user capacity region of the white Gaussian channel which is equal to the Cover–Wyner pentagon, whereas if the assigned waveforms are different (e.g., code division multiple access), the symbol-asynchronous capacity region is no longer a pentagon.

I. Introduction

THE MAIN GOAL of the information-theoretic study of the multiple-access channel is to find its capacity region, i.e., the set of information rates at which simultaneous reliable communication of the messages of each user is possible. This problem was solved in the pioneering work of Ahlswede [1], [2] on the two-user discrete memoryless channel; later, an explicit expression for the capacity region of the Gaussian memoryless discrete-time multiple-access channel was given by Cover [3] and Wyner [4]. These and most of the subsequent results on the subject assumed so-called frame (or block) synchronism among the users in the sense that the beginnings of the codewords of each user were guaranteed to coincide at the receiver. It has been shown by Poltyrev [5] and, independently, by Hui and Humblet [6] that the only effect of frame asynchronism on the discrete memoryless multiple-access channel is the removal of the convex hull operation from the expression for the capacity region. It was recently shown [7] that if the multiple-access channel has memory, frame asynchronism may drastically reduce the capacity region and, in particular, the maximum achievable rate sum. At any rate, in many practical situations it is perfectly reasonable to assume that this type of synchronism is achievable with a modicum of channel feedback or cooperation among transmitters.

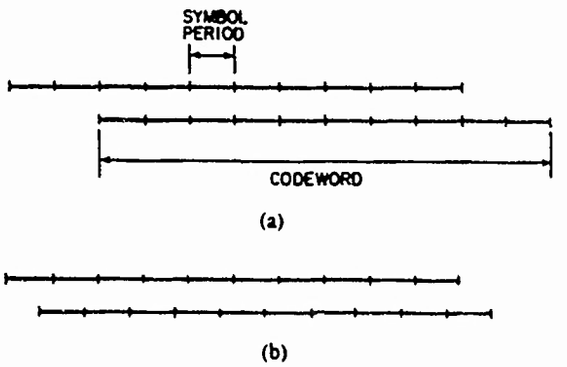

Fig. 1. (a) Frame-asynchronous symbol-synchronous two-user channel. (b) Frame-synchronous symbol-asynchronous two-user channel.

The type of synchronism that is difficult to achieve in many practical situations (owing to the much smaller time scale involved) is symbol synchronism. This issue arises in continuous-time channels where each codeword symbol modulates a signal waveform of finite duration, as is the case in most conventional digital communication systems. In these systems, user k transmits a codeword $(b_k(1),\cdots,b_k(n)) \in A_k^n$ by sending the signal

$$\sum_{i=1}^{n} s_k(t-iT; b_k(i))$$

where the waveforms $\{s_k(t; b),\ b \in A_k\}$ vanish outside the interval $[0,T]$ and constitute the fixed signaling alphabet of user k, which is known to all transmitters and to the receiver. If the symbol epochs of the signals transmitted by the users are not aligned at the receiver, then the channel is symbol asynchronous (Fig. 1). For a channel with two senders and one receiver, assuming frame synchronism and an additive white Gaussian noise channel model, we can write the channel output as

$$y(t) = \sum_{i=1}^{n} s_1(t-iT-\tau_1; b_1(i)) + \sum_{i=1}^{n} s_2(t-iT-\tau_2; b_2(i)) + n(t) \quad (1.1)$$

Manuscript received June 6, 1987; revised October 10, 1988. This work was supported in part by the Office of Naval Research under Contract N00014-87-K-0054. This paper was presented in part at the IEEE International Symposium on Information Theory, Ann Arbor, MI, October 1986, and in part at the IEEE Workshop on Information Theory, Bellagio, Italy, June, 1987.

The author is with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544.

IEEE Log Number 8929032.

0018-9448/89/0700-0733$01.00 ©1989 IEEE

where the delays or offsets $\tau_1 \in [0, T)$, $\tau_2 \in [0, T)$ account for the symbol asynchronism between the users and are known to the receiver (because it acquires the timing of each of the received signals to decode reliably each of the transmitted messages) and unknown to the transmitters.

While the derivation of coding strategies for symbol-asynchronous channels has been addressed before [8], it appears that no results on the capacity region of the multiple-access channel are available when symbol synchronism is not assumed. In this paper we find the capacity region of a fairly general symbol-asynchronous Gaussian multiple-access channel in which user k modulates linearly a fixed *signature waveform* $s_k(t)$, i.e., $s_k(t; b) = bs_k(t)$. This encompasses many interesting channels in applications, such as direct-sequence spread-spectrum code-division multiple-access channels (CDMA) wherein each transmitter is assigned a distinct signature waveform which is used to modulate information simultaneously and independently of the other transmitters.[1] We focus our attention on energy-limited channels where $A_k = \mathbf{R}$, and each codeword of user k is constrained to satisfy

$$\frac{1}{n}\sum_{i=1}^{n} b_k^2(i) \le w_k, \qquad k = 1, 2. \tag{1.2}$$

The methods employed in this paper can be used to solve the case where the A_k are finite alphabets; however, in this case, as in the single-user discrete-time Gaussian channel with finite alphabets or amplitude constraints, no explicit expressions for capacity can be obtained.

If the transmitters are assigned identical signature waveforms and are symbol synchronous, i.e., $\tau_1 = \tau_2$, then it is easy to see that the channel is equivalent to the standard one-dimensional discrete-time Gaussian multiple-access channel, and therefore, its capacity region is given by the Cover–Wyner pentagon: each individual rate is constrained not to exceed single-user capacity and the sum of the rates cannot exceed the capacity of a single-user channel whose signal-to-noise ratio is the sum of the signal-to-noise ratios of both users. In this paper it is shown that the same result holds even if the users are not symbol synchronous. However, that is no longer true when the transmitters are assigned different signature waveforms. Then the symbol-asynchronous capacity region is no longer a pentagon and depends not only on the respective signal-to-noise ratios, but also on the similarity between the signature waveforms quantified by their cross correlations. In some applications it may be of interest to use the capacity region found in this paper for any arbitrary choice of signature waveforms as a basis for optimum signal design (i.e., to find the elements that achieve the boundary of the union of capacity regions over a certain set of signature waveforms) under a variety of specific constraints on the set of feasible signals, such as direct-sequence waveforms with a maximum number of chips-per-symbol or signals approximately bandlimited to a specified bandwidth. However, it is worth noting that in many practical applications the choice of signature waveforms is dictated by considerations such as jamming resistance and the use of specific waveforms selected from families of pseudonoise sequences with favorable cross-correlation properties (such as Gold sequences or maximal-length shift-register sequences).

The first step in the derivation of the capacity of the symbol-asynchronous Gaussian channel is to obtain an equivalent channel model with discrete-time outputs. This is the purpose of Section II, where an equivalent discrete-time Gaussian channel parametrized by the signal cross correlations is derived. The main feature introduced by the lack of symbol synchronism is that the channel has memory. This is due to the overlap of each symbol transmitted by a user with two consecutive symbols transmitted by the other user. The capacity of discrete-time multiple-access channels with finite memory was obtained in [7] with and without frame synchronization. Those results are used in Section III to obtain the capacity region of the symbol-asynchronous Gaussian multiple-access channel, which turns out to be independent of whether or not the channel is frame synchronous. Since the relative offset $\tau_2 - \tau_1$ between the received signals is not known to the transmitters, we must deal with a compound multiple-access channel where the encoders only know that the actual channel belongs to an uncertainty set parametrized by the relative offset. For the sake of clarity of exposition we deal first with the case where the relative offset is known to all parties (i.e., the uncertainty set is a singleton), and then we use those results to find the sought-after capacity region of the compound channel. Finally, in Section IV we consider an alternative representation of the capacity region which results in a particularly compact characterization of the fundamental limits of the multiple-access channel in the region of high signal-to-noise ratios.

II. Channel Model

The goal of this section is to obtain a channel with discrete-time outputs whose capacity is the same as that of the channel with continuous-time output

$$y(t) = \sum_{i=1}^{n} b_1(i)s_1(t - iT - \tau_1) + \sum_{i=1}^{n} b_2(i)s_2(t - iT - \tau_2) + n(t) \tag{2.1}$$

where $n(t)$ is white Gaussian noise with power spectral density equal to σ^2. This goal is achieved by considering the projection of the observation process $\{y(t)\}$ along the direction of the unit energy signals $\{s_1(t)\}$ and $\{s_2(t)\}$ and their T-shifts:

$$y_k(i) = \int_{iT+\tau_k}^{(i+1)T+\tau_k} y(t)s_k(t - iT - \tau_k)\,dt, \qquad k = 1, 2. \tag{2.2}$$

[1] Most capacity analyses of the CDMA channel have focused on single-user receivers and approximated the multiple-access interference by a white Gaussian process [9]–[11], thereby providing limited insight into the fundamental limits of that channel.

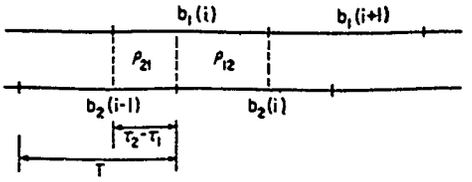

Fig. 2. Symbol periods and cross correlations.

It is possible to obtain an expression for $\{y_1(i)\}_{i=1}^n$ and $\{y_2(i)\}_{i=1}^n$ in terms of the transmitted codewords $\{b_1(i)\}_{i=1}^n$ and $\{b_2(i)\}_{i=1}^n$ by substituting (2.1) into (2.2) and by defining the cross correlations between the assigned signature waveforms $\{s_1(t)\}$ and $\{s_2(t)\}$ as (assuming without loss of generality that $\tau_1 \le \tau_2$ (Fig. 2))

$$\rho_{12} = \int_0^T s_1(t) s_2(t+\tau_1-\tau_2)\, dt \tag{2.3a}$$

$$\rho_{21} = \int_0^T s_1(t) s_2(t+T+\tau_1-\tau_2)\, dt. \tag{2.3b}$$

It follows easily that

$$\begin{bmatrix} y_1(i) \\ y_2(i) \end{bmatrix} = \begin{bmatrix} 0 & \rho_{21} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} b_1(i-1) \\ b_2(i-1) \end{bmatrix} + \begin{bmatrix} 1 & \rho_{12} \\ \rho_{12} & 1 \end{bmatrix} \begin{bmatrix} b_1(i) \\ b_2(i) \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ \rho_{21} & 0 \end{bmatrix} \begin{bmatrix} b_1(i+1) \\ b_2(i+1) \end{bmatrix} + \begin{bmatrix} n_1(i) \\ n_2(i) \end{bmatrix} \tag{2.4}$$

for $1 \le i \le n$ (with $b_k(0) = b_k(n+1) = 0$, $k = 1,2$); the discrete-time random process $\{[n_1(i)\ n_2(i)]^T\}$ is Gaussian with zero mean and covariance matrix:

$$E\left[\begin{bmatrix} n_1(i) \\ n_2(i) \end{bmatrix} [n_1(j)\ n_2(j)]\right] = \sigma^2 H(i-j)$$

where $H(i) = 0$ if $|i| > 1$, and $H(1)$, $H(0)$, and $H(-1)$ are the matrices appearing in (2.4), i.e.,

$$H(0) = \begin{bmatrix} 1 & \rho_{12} \\ \rho_{12} & 1 \end{bmatrix} \qquad H(1) = H^T(-1) = \begin{bmatrix} 0 & \rho_{21} \\ 0 & 0 \end{bmatrix}.$$

Since the receiver knows the assigned waveforms $\{s_1(t)\}$ and $\{s_2(t)\}$ as well as the symbol epochs $\{iT+\tau_1\}$ and $\{iT+\tau_2\}$, it can compute $\{y_1(i)\}_{i=1}^n$ and $\{y_2(i)\}_{i=1}^n$ by passing the observations through two matched filters for signals $\{s_1(t)\}$ and $\{s_2(t)\}$, respectively. The key observation is that this operation does not destroy any information that is valuable in deciding which messages were transmitted. This is because the likelihood function (i.e., the conditional expectation of the Radon–Nikodym derivative between the measure induced by the observations and Wiener measure given that $\{b_1(i)\}_{i=1}^n$ and $\{b_2(i)\}_{i=1}^n$ are the transmitted codewords) is equal to a constant times (e.g., [12])

$$\exp\left\{-\int \left[y(t) - \sum_{k=1}^{2} \sum_{i=1}^{n} b_k(i) s_k(t-iT-\tau_k)\right]^2 dt\right\}$$

which can be factored into

$$h_1(\{y(t)\}) h_2(\{b_1(i)\}, \{b_2(i)\}, \{y_1(i)\}, \{y_2(i)\});$$

hence because of the factorization theorem [13], $\{y_1(i)\}_{i=1}^n$ and $\{y_2(i)\}_{i=1}^n$ are *sufficient statistics* for the transmitted messages. This implies that the channel output $\{y(t)\}$ enters in the computation of the posterior probability of each message only through $\{y_1(i)\}_{i=1}^n$ and $\{y_2(i)\}_{i=1}^n$. Thus, no matter which codebooks are chosen by the transmitters, the probability that the maximum *a posteriori* decoder selects the true transmitted message remains the same if instead of working with the original continuous-time observations $\{y(t)\}$ the decoder is constrained to work with the discrete-time sequences $\{y_1(i)\}_{i=1}^n$ and $\{y_2(i)\}_{i=1}^n$. Therefore, if a rate pair is ϵ-achievable for the multiple-access channel (2.1), it is also ϵ-achievable for the multiple-access channel (2.4), and hence the capacity regions of both channels coincide. In this respect, notice for future use that if $t(W)$ is a sufficient statistic for Z, then the data-processing inequality is satisfied with equality because Z and W are conditionally independent given $t(W)$. Therefore,

$$\begin{aligned} I(Z; t(W)) &= I(Z; W, t(W)) - I(Z; W|t(W)) \\ &= I(Z; W) + I(Z; t(W)|W) \\ &= I(Z; W). \end{aligned} \tag{2.5}$$

Note that even though channel (2.4) has two output sequences, it is a multiple-access channel rather than an interference channel because both outputs are available to the multiuser receiver. Channel (2.4) is parametrized by the cross correlations ρ_{12} and ρ_{21}, which depend on the relative offset $\tau_2 - \tau_1$ and, therefore, in general, are unknown to the transmitters. Consequently, it is necessary to analyze a *compound* multiple-access channel where the transmitters only know that (ρ_{12}, ρ_{21}) belongs to an uncertainty set determined by $\{s_1(t)\}$ and $\{s_2(t)\}$.

The main characteristic of the discrete-time multiple-access channel in (2.4) is that it has memory because the noise sequence is correlated and each output value depends on three input symbols, while each of these symbols affects two consecutive output vectors (cf. Fig. 2). It is possible to obtain an equivalent multiple-access channel (Appendix I) whose noise process is independent at the expense of an enlarged set of observables. The advantage of the latter discrete-time model is that it is possible to invoke coding theorems for channels where the outputs are conditionally independent given the inputs [7].

If either $\rho_{12} = 0$ or $\rho_{21} = 0$, then the channel becomes memoryless because in that case the users are in effect symbol synchronous. For example, if the users are assigned the same signal and the channel is symbol synchronous, both outputs in (2.4) coincide and are equal to

$$y(i) = b_1(i) + b_2(i) + n(i) \tag{2.6}$$

where $\{n(i)\}$ is a Gaussian independent sequence. Then the channel is the conventional scalar discrete-time Gaussian channel, whose capacity region, subject to the energy

onstraints in (1.2), is the Cover–Wyner region:

$$C = \left\{(R_1, R_2): 0 \le R_1 \le \frac{1}{2}\log\left[1+\frac{w_1}{\sigma^2}\right]\right.$$
$$0 \le R_2 \le \frac{1}{2}\log\left[1+\frac{w_2}{\sigma^2}\right]$$
$$\left. R_1 + R_2 \le \frac{1}{2}\log\left[1+\frac{w_1}{\sigma^2}+\frac{w_2}{\sigma^2}\right]\right\}. \tag{2.7}$$

If the assigned signals are not equal but the users remain symbol synchronous, then (2.4) reduces to the memoryless multiple-access channel

$$\begin{bmatrix} y_1(i) \\ y_2(i) \end{bmatrix} = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\begin{bmatrix} b_1(i) \\ b_2(i) \end{bmatrix} + \begin{bmatrix} n_1(i) \\ n_2(i) \end{bmatrix} \tag{2.8}$$

where $\{[n_1(i)n_2(i)]^T\}$ is an independent Gaussian process with $E[n_k^2(i)] = \sigma^2$ and $E[n_1(i)n_2(i)] = \sigma^2\rho$, and

$$\rho = \int_0^T s_1(t)s_2(t)\,dt.$$

In this case, the Cover–Wyner region can be easily generalized (Section III) thanks to the lack of memory when the users are symbol synchronous.

III. Capacity Region

Before we obtain the capacity region of the symbol-asynchronous Gaussian multiple-access channel, we will generalize the Cover–Wyner region (2.7) to the symbol-synchronous channel where both users are not necessarily assigned the same waveform. To this end, according to (2.8) we need to find the convex closure over independent random variables X_1 and X_2 such that $E[X_1^2] \le w_1$ and $E[X_2^2] \le w_2$ of the union of the pentagons $\{0 \le R_1 \le I(X_1; Y|X_2), 0 \le R_2 \le I(X_2; Y|X_1), \quad R_1 + R_2 \le I(X_1, X_2; Y)\}$, with the output Y given by

$$Y = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} + \begin{bmatrix} N_1 \\ N_2 \end{bmatrix} \tag{3.1}$$

where N_1 and N_2 are jointly Gaussian with zero mean, $E[N_k^2] = \sigma^2$, and $E[N_1N_2] = \rho\sigma^2$. The case $|\rho| = 1$ results in the region (2.7); we will therefore assume $|\rho| < 1$. Since X_1 and X_2 are independent random variables, the covariance of Y is equal to[2]

$$\text{cov}(Y) = \begin{vmatrix} 1 & \rho \\ \rho & 1 \end{vmatrix}\left|\sigma^2 I_2 + \begin{bmatrix} \text{var}(X_1) & 0 \\ 0 & \text{var}(X_2) \end{bmatrix}\begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right|, \tag{3.2}$$

[2] I_n denotes the $n \times n$ identity matrix.

and we can upper-bound the mutual informations

$$I(X_1, X_2; Y)$$
$$\le \frac{1}{2}\log\det[\text{cov}(Y)] - \frac{1}{2}\log\det\left[\sigma^2\begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right]$$
$$= \frac{1}{2}\log\det\left[I_2 + \frac{1}{\sigma^2}\begin{bmatrix} \text{var}(X_1) & 0 \\ 0 & \text{var}(X_2) \end{bmatrix}\begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right]$$
$$= \frac{1}{2}\log\left[1 + \frac{\text{var}(X_1)}{\sigma^2} + \frac{\text{var}(X_2)}{\sigma^2} + \frac{\text{var}(X_1)}{\sigma^2}\frac{\text{var}(X_2)}{\sigma^2}(1-\rho^2)\right] \tag{3.3}$$

$$I(X_1; Y|X_2) \le \frac{1}{2}\log\det\left[I_2 + \frac{1}{\sigma^2}\begin{bmatrix} \text{var}(X_1) & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right]$$
$$= \frac{1}{2}\log\left(1 + \frac{\text{var}(X_1)}{\sigma^2}\right) \tag{3.4}$$

and similarly,

$$I(X_2; Y|X_1) \le \frac{1}{2}\log\left(1 + \frac{\text{var}(X_2)}{\sigma^2}\right) \tag{3.5}$$

with equality in (3.3), (3.4), and (3.5) if X_1 and X_2 are Gaussian. Furthermore, all three rate constraints are simultaneously maximized by letting X_1 and X_2 attain the maximum allowable variances, i.e., w_1 and w_2, respectively. Hence the capacity region is equal to the pentagon

$$C = \left\{(R_1, R_2): 0 \le R_1 \le \frac{1}{2}\log\left(1+\frac{w_1}{\sigma^2}\right)\right.$$
$$0 \le R_2 \le \frac{1}{2}\log\left(1+\frac{w_2}{\sigma^2}\right)$$
$$\left. R_1 + R_2 \le \frac{1}{2}\log\left[1+\frac{w_1}{\sigma^2}+\frac{w_2}{\sigma^2}+\frac{w_1w_2}{\sigma^4}(1-\rho^2)\right]\right\} \tag{3.6}$$

which differs from (2.7) in that the maximum rate sum is no longer the capacity of a single-user channel whose signal-to-noise ratio is equal to $(w_1 + w_2)/\sigma^2$. Notice that when $\{s_1(t)\}$ and $\{s_2(t)\}$ are orthogonal ($\rho = 0$), then, effectively, both users transmit in separate noninterfering channels and can send information at single-user rates. The K-user capacity region of the symbol-synchronous Gaussian channel can be found in [14].

Before we state and prove the formula for the capacity of the symbol-asynchronous Gaussian multiple-access channel, we will motivate the expression of the capacity region by finding the mutual information rates in channel (2.4) when the inputs are stationary Gaussian processes with power spectral densities $\{S_1(\omega), \omega \in [-\pi, \pi]\}$ and $\{S_2(\omega), \omega \in [-\pi, \pi]\}$. Channel (2.4) is a two-input two-

›utput dynamic linear time-invariant system whose out-›uts are embedded in colored stationary Gaussian noise. If he inputs are stationary Gaussian processes, then the nutual information rates can be written as the difference ›etween the differential entropy rates of the output with :nd without each of the input processes. Consequently, all s needed is an expression for the differential entropy rate ›f a stationary vector Gaussian discrete-time process. In he scalar case the differential entropy rate of a Gaussian iscrete-time process whose power spectral density is $S(\omega)$ s equal to [15, p. 542]

$$h(S) = \frac{1}{2}\log(2\pi e L(S)) \tag{3.7}$$

vnere $L(S)$ is the *geometric mean* of $S(\omega)$, i.e.,

$$L(S) = \exp\frac{1}{2\pi}\int_{-\pi}^{\pi}\log S(\omega)\,d\omega. \tag{3.8}$$

This follows because the differential entropy of a Gaussian i-vector with covariance matrix Σ_n is $n/2)\log(2\pi e(\det\Sigma_n)^{1/n})$ and according to the Toeplitz istribution theorem [16], $\lim_{n\to\infty}(\det\Sigma_n)^{1/n}$ coincides vith the geometric mean of the Fourier transform of the ovariance sequence. What we need for our purposes is a eneraiization of this result to vector random processes, .e., we need to find $\lim_{n\to\infty}(\det\Sigma_n)^{1/n}$ when Σ_n is an i-block Toeplitz matrix whose elements are 2×2 covari-nce matrices $R(i-j) \triangleq \Sigma_n(i,j)$. A solution to this prob-em can be found in [17] where it is shown that if the ›ower spectral density matrix $M(\omega) = \sum_{n=-\infty}^{\infty} e^{-j\omega n}R(n)$ is ontinuous and positive definite in $[-\pi,\pi]$,[3] then the oregoing limit is equal to the geometric mean of the eterminant of $M(\omega)$.

Now the output of channel (2.4) is a zero-mean vector aussian process with power spectral density matrix given

Therefore, the mutual information rate between the output and the inputs is equal to

$$\begin{aligned}\lim_{n\to\infty}\frac{1}{n}I(X_1^n, X_2^n; Y^n) &= h(\Xi) - h(\sigma^2 T)\\ &= \frac{1}{4\pi}\int_{-\pi}^{\pi}\log\det\left[I_2 + \frac{1}{\sigma^2}\begin{bmatrix} S_1(\omega) & 0\\ 0 & S_2(\omega)\end{bmatrix}T(\omega)\right]d\omega\\ &= \frac{1}{4\pi}\int_{-\pi}^{\pi}\log\left(1+\frac{S_1(\omega)}{\sigma^2}+\frac{S_2(\omega)}{\sigma^2}\right.\\ &\quad\left.+\frac{S_1(\omega)S_2(\omega)}{\sigma^4}\left[1-\rho^2(\omega)\right]\right)d\omega\end{aligned} \tag{3.9}$$

where

$$\rho^2(\omega) = |\rho_{12} + \rho_{21}e^{j\omega}|^2 = \rho_{12}^2 + \rho_{21}^2 + 2\rho_{21}\rho_{12}\cos\omega. \tag{3.10}$$

Similarly, setting $S_2(\omega) = 0$ and $S_1(\omega) = 0$, respectively, in 3.9) we get

$$\lim_{n\to\infty}\frac{1}{n}I(X_1^n; Y^n|X_2^n) = \frac{1}{4\pi}\int_{-\pi}^{\pi}\log\left(1+\frac{S_1(\omega)}{\sigma^2}\right)d\omega \tag{3.11}$$

$$\lim_{n\to\infty}\frac{1}{n}I(X_2^n; Y^n|X_1^n) = \frac{1}{4\pi}\int_{-\pi}^{\pi}\log\left(1+\frac{S_2(\omega)}{\sigma^2}\right)d\omega. \tag{3.12}$$

As mentioned in Section I, we will find the capacity region of the asynchronous channel first in the case where the transmitters know the offset, and hence the cross correlations between their signals, and then in the case where they do not.

Theorem 1: The capacity region of the energy-constrained asynchronous Gaussian multiple-access channel when the transmitters know their mutual offset is given by

$$C = \bigcup_{\substack{S_i(\omega)\ge 0,\ \omega\in[-\pi,\pi]\\ \frac{1}{2\pi}\int_{-\pi}^{\pi}S_i(\omega)\,d\omega = w_i\\ i=1,2}} I(R_1,R_2),\ 0\le R_1\le\frac{1}{4\pi}\int_{-\pi}^{\pi}\log\left(1+\frac{S_1(\omega)}{\sigma^2}\right)d\omega,\ 0\le R_2\le\frac{1}{4\pi}\int_{-\pi}^{\pi}\log\left(1+\frac{S_2(\omega)}{\sigma^2}\right)d\omega$$

$$R_1 + R_2 \le \frac{1}{4\pi}\int_{-\pi}^{\pi}\log\left(1+\frac{S_1(\omega)}{\sigma^2}+\frac{S_2(\omega)}{\sigma^2}+\frac{S_1(\omega)S_2(\omega)}{\sigma^4}\left[1-\rho_{12}^2-\rho_{21}^2-2\rho_{12}\rho_{21}\cos\omega\right]\right)d\omega\Big\}. \tag{3.13}$$

›v

$$\Xi(\omega) = T(\omega)\begin{bmatrix} S_1(\omega) & 0\\ 0 & S_2(\omega)\end{bmatrix}T(\omega) + \sigma^2 T(\omega)$$

$$T(\omega) = \begin{bmatrix} 1 & \rho_{12}+\rho_{21}e^{-j\omega}\\ \rho_{12}+\rho_{21}e^{j\omega} & 1\end{bmatrix}.$$

[3] In the present case the power spectral density of the output vector process is indeed continuous, but in problems with heavily correlated waveforms it may fail to be nonsingular at particular frequencies. However, the capacity region is derived later in this section without imposing any of those assumptions.

Proof: It is shown in [7, Theorem 3] that the capacity region of the frame-synchronous discrete-time multiple-access channel with finite memory where the outputs depend on several consecutive input symbols and are conditionally independent given the inputs is equal to

$$C = \text{closure}\left(\liminf_{n\to\infty}\frac{1}{n}C_n\right) \tag{3.14}$$

where C_n is the following achievable region for the n-block memoryless multiple-access channel whose input symbols correspond to n consecutive channel uses of the original

channel with memory:

$$C_n = \bigcup_{X_1^n, X_2^n} \{(R_1^n, R_2^n): 0 \le R_1^n \le I(X_1^n; Y^n | X_2^n)$$
$$0 \le R_2^n \le I(X_2^n; Y^n | X_1^n)$$
$$R_1^n + R_2^n \le I(X_1^n, X_2^n; Y^n)\} \quad (3.15)$$

where the union is over the independent random variables X_1^n and X_2^n satisfying in the present case $E[X_k^{nT} X_k^n] \le w_k$, $k = 1,2$.

The aforementioned class of multiple-access channels with finite memory includes as a special case the discrete-time channel in (I.3) whose noise sequence is independent. It does not encompass channel (2.4) directly because the noise sequence therein is dependent. However, since the observables of (2.4) are sufficient statistics for the inputs and are deterministic transformations of the redundant observables in (I.3), not only their capacities coincide but, according to (2.5), the mutual informations arising in the achievable regions, C_n, of the respective induced n-block memoryless channels are also equal. Therefore, it is enough to show that the closed set in (3.13) is equal to $\lim_{n\to\infty}(1/n)C_n$, where C_n is the achievable region in (3.15) for the n-block memoryless multiple-access channel induced by (2.4). To this end it is easy to check using (2.4) that the n-block multiple-access channel can be written as

$$Y^n = \begin{bmatrix} Y_1(1) \\ Y_2(1) \\ Y_1(2) \\ \vdots \\ Y_1(n) \\ Y_2(n) \end{bmatrix} = \begin{bmatrix} 1 & \rho_{12} & & & \\ \rho_{12} & 1 & \rho_{21} & & \\ & \rho_{21} & 1 & \rho_{12} & \\ & & \ddots & \ddots & \ddots \\ & & & \rho_{21} & 1 & \rho_{12} \\ & & & & \rho_{12} & 1 \end{bmatrix} \begin{bmatrix} X_1(1) \\ X_2(1) \\ X_1(2) \\ \vdots \\ X_1(n) \\ X_2(n) \end{bmatrix} + \begin{bmatrix} N_1(1) \\ N_2(1) \\ N_1(2) \\ \vdots \\ N_1(n) \\ N_2(n) \end{bmatrix} \quad (3.16)$$

which is depicted in Fig. 3 and where according to (2.4), the noise vector is Gaussian with zero mean and covariance matrix $\sigma^2 \boldsymbol{R}$, where $\boldsymbol{R}$ is the block diagonal $2n \times 2n$ cross correlation matrix multiplying the input vector in (3.15). This is a positive-definite matrix because $x^T \boldsymbol{R} x$ is equal to the energy of $\sum_{i=1}^{n}\sum_{k=1}^{2} x_k(i) s_k(t - \tau_k - iT)$ which is guaranteed to be nonzero if $x \neq 0$, $\rho_{12} \neq 0$, and $\rho_{21} \neq 0$. Throughout this proof we assume that $\rho_{12} \neq 0$ and $\rho_{21} \neq 0$; otherwise, the channel is equivalent to a symbol-synchronous channel and the capacity region is given by (3.6) (which coincides with (3.13) because if $\rho_{12}\rho_{21} = 0$, then the three rate constraints therein are maximized by white spectra).

The output covariance matrix is equal to

$$\operatorname{cov}(Y^n) = \boldsymbol{R}\left[\sigma^2 I_{2n} + E[X^n X^{nT}]\boldsymbol{R}\right] \quad (3.17)$$

where[4]

$$X^n = [X_1(1), X_2(1), X_1(2), \cdots X_1(n), X_2(n)]^T$$
$$= X_1^n \otimes \begin{bmatrix} 1 \\ 0 \end{bmatrix} + X_2^n \otimes \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \quad (3.18)$$

[4] $A \otimes B$ denotes the Kronecker product of the matrices A and B.

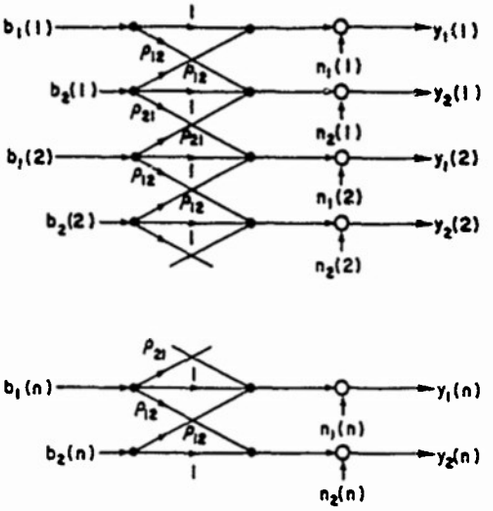

Fig. 3. n-block memoryless two-user channel.

As in (3.3), (3.4), and (3.5) we can upper-bound the mutual informations by

$$I(X_1^n, X_2^n; Y^n) \le \frac{1}{2}\log\det[\operatorname{cov}(Y^n)] - \frac{1}{2}\log\det[\sigma^2 \boldsymbol{R}]$$
$$= \frac{1}{2}\log\det\left[I_{2n} + \sigma^{-2} E[X^n X^{nT}]\boldsymbol{R}\right] \quad (3.19)$$

and

$$I(X_1^n; Y^n | X_2^n) \le \frac{1}{2}\log\det\left[I_n + \sigma^{-2}\Sigma_1\right] \quad (3.20a)$$

$$I(X_2^n; Y^n | X_1^n) \le \frac{1}{2}\log\det\left[I_n + \sigma^{-2}\Sigma_2\right] \quad (3.20b)$$

where $\Sigma_k = \operatorname{cov}(X_k^n)$, $k = 1,2$, and equality holds in (3.19) and (3.20) if X_1^n and X_2^n are Gaussian. The following identity whose proof is in Appendix II gives an explicit expression of (3.19) as a function of Σ_1, Σ_2, ρ_{12}, and ρ_{21}.

Lemma 1: The following identity holds

$$\det\left[I_{2n} + \sigma^{-2} E[X^n X^{nT}]\boldsymbol{R}\right]$$
$$= \det\left[I_{2n} + \frac{1}{\sigma^2}\begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix}\begin{bmatrix} I_n & S^T \\ S & I_n \end{bmatrix}\right] \quad (3.21)$$

where

$$S = \rho_{12} I_n + \rho_{21}\begin{bmatrix} 0 & 1 & 0 & & & \\ 0 & 0 & 1 & & & \\ & 0 & 0 & \vdots & 0 & \\ & & & & 1 & 0 \\ & & & & 0 & 1 \\ & & & & 0 & 0 \end{bmatrix}. \quad (3.22)$$

Therefore, (3.15) reduces to the following union over all trace-constrained nonnegative-definite $n \times n$ matrices:

$$C_n = \bigcup_{\substack{\Sigma_k \geq 0 \\ \frac{1}{n}\mathrm{tr}\Sigma_k \leq w_k \\ k=1,2}} \left\{ (R_1, R_2), 0 \leq R_1 \leq \frac{1}{2}\log\det\left[I_n + \sigma^{-2}\Sigma_1\right] \right.$$

$$0 \leq R_2 \leq \frac{1}{2}\log\det\left[I_n + \sigma^{-2}\Sigma_2\right]$$

$$\left. R_1 + R_2 \leq \frac{1}{2}\log\det\left[I_{2n} + \frac{1}{\sigma^2}\begin{bmatrix}\Sigma_1 & 0 \\ 0 & \Sigma_2\end{bmatrix}\begin{bmatrix}I_n & S^T \\ S & I_n\end{bmatrix}\right] \right\}. \tag{3.23}$$

Region C_n is a convex set because each of the three rate constraints in (3.21) is a concave function of (Σ_1, Σ_2). This is a consequence of the fact that $\log\det[A]$ is concave in A if A is a positive-definite matrix [18, p. 125]. Note that even though the determinant appearing in the rate-sum constraint is not that of a symmetric matrix, it is equal to the determinant of the positive-definite matrix

$$I_{2n} + \frac{1}{\sigma^2}Q\begin{bmatrix}\Sigma_1 & 0 \\ 0 & \Sigma_2\end{bmatrix}Q^T, \quad \text{where} \quad Q^TQ = \begin{bmatrix}I_n & S^T \\ S & I\end{bmatrix}.$$

There is no covariance pair (Σ_1, Σ_2) that maximizes all three rate constraints in (3.23) simultaneously. This is in contrast to the symbol-synchronous channel where we saw that the mutual informations in (3.3)–(3.5) are maximized simultaneously by a pair of input distributions, thus resulting in a capacity region which is equal to a pentagon. Nevertheless, we will be able to show that there is a set of optimum eigenvectors for each user in the sense that it is enough to take the union in (3.23) only over the subset of covariance matrices having those eigenvectors, thereby effectively reducing the union to one over diagonal matrices. To prove this, the first step is to apply the singular-value decomposition theorem to the matrix S defined in (3.22). According to this result [19, p. 192], we can write

$$S = UDV^* \tag{3.24}$$

where U and V are orthogonal matrices (of the eigenvectors of SS^T and S^TS, respectively) and D is a diagonal matrix of the singular values $\{d_i\}_{i=1}^n$ of S, i.e., the nonnegative square roots of the eigenvalues of the nonnegative-definite Jacobi matrix

$$S^TS = \begin{bmatrix} \rho_{12}^2 & \rho_{12}\rho_{21} & & & & \\ \rho_{12}\rho_{21} & \rho_{12}^2+\rho_{21}^2 & \rho_{12}\rho_{21} & & & \\ & \rho_{12}\rho_{21} & \rho_{12}^2+\rho_{21}^2 & \cdot & & \\ & & & \cdot & \rho_{12}\rho_{21} & \\ & & & \cdot & \rho_{12}^2+\rho_{21}^2 & \rho_{12}\rho_{21} \\ & & & & \rho_{12}\rho_{21} & \rho_{12}^2+\rho_{21}^2 \end{bmatrix}. \tag{3.25}$$

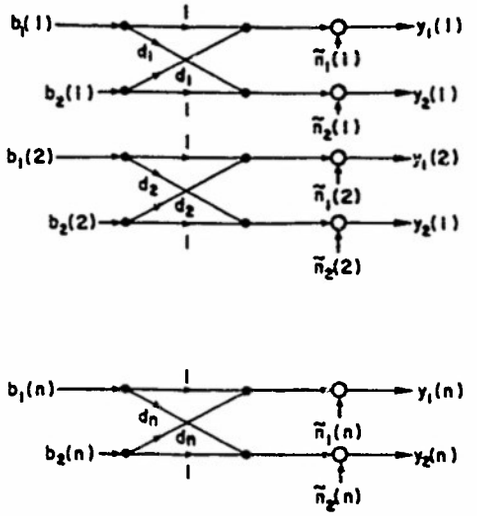

Fig. 4. Decoupled n-block memoryless two-user channel.

Now, using the orthogonality of U and V, we can express the determinant in the rate-sum constraint in (3.23) as

$$\begin{aligned}\det\begin{bmatrix}V^* & 0 \\ 0 & U^*\end{bmatrix}\det\left[I_{2n} + \frac{1}{\sigma^2}\begin{bmatrix}\Sigma_1 & 0 \\ 0 & \Sigma_2\end{bmatrix}\begin{bmatrix}I_n & S^T \\ S & I_n\end{bmatrix}\right]\det\begin{bmatrix}V & 0 \\ 0 & U\end{bmatrix} \\ = \det\left[I_{2n} + \frac{1}{\sigma^2}\begin{bmatrix}V^*\Sigma_1V & V^*\Sigma_1VD \\ U^*\Sigma_2UD & U^*\Sigma_2U\end{bmatrix}\right] \\ = \det\left[I_{2n} + \frac{1}{\sigma^2}\begin{bmatrix}\Lambda_1 & 0 \\ 0 & \Lambda_2\end{bmatrix}\begin{bmatrix}I_n & D \\ D & I_n\end{bmatrix}\right]\end{aligned} \tag{3.26}$$

where we have set $\Lambda_1 = V^*\Sigma_1V$ and $\Lambda_2 = U^T\Sigma_2U$. Since $\mathrm{tr}(\Lambda_k) = \mathrm{tr}(\Sigma_k)$, $\Sigma_k \geq 0$ if and only if $\Lambda_k \geq 0$, and

$$\det\left[I_n + \sigma^{-2}\Lambda_k\right] = \det\left[I_n + \sigma^{-2}\Sigma_k\right], \tag{3.27}$$

the region in (3.23) is equal to

$$C_n = \bigcup_{\substack{\Lambda_k \geq 0 \\ \mathrm{tr}\Lambda_k \leq nw_k \\ k=1,2}} \left\{ (R_1, R_2), 0 \leq R_1 \leq \frac{1}{2}\log\det\left[I_n + \sigma^{-2}\Lambda_1\right] \right.$$

$$\cdot 0 \leq R_2 \leq \frac{1}{2}\log\det\left[I_n + \sigma^{-2}\Lambda_2\right]$$

$$\left. \cdot R_1 + R_2 \leq \frac{1}{2}\log\det\left[I_{2n} + \frac{1}{\sigma^2}\begin{bmatrix}\Lambda_1 & 0 \\ 0 & \Lambda_2\end{bmatrix}\begin{bmatrix}I_n & D \\ D & I_n\end{bmatrix}\right] \right\}. \tag{3.28}$$

Thus in effect the singular-value decomposition theorem has allowed us to substitute the matrix S in (3.23) by the diagonal matrix D. This is advantageous because the set in (3.28) is actually the capacity region of the two-user Gaussian memoryless channel shown in Fig. 4. This channel differs from the one in Fig. 3 in that the inputs corresponding to different coordinates do not interfere, and the noise covariance matrix is

$$E\left[\begin{bmatrix}\tilde{n}_1(i)\\ \tilde{n}_2(j)\end{bmatrix}[\tilde{n}_1(i)\tilde{n}_2(j)]\right]=\sigma^2\delta_{ij}\begin{bmatrix}1 & d_i\\ d_i & 1\end{bmatrix}. \quad (3.29)$$

Therefore, the singular-value decomposition of S effectively decouples the original channel in (3.16) into independent 2×2 multiple-access subchannels. The capacity region of this decoupled channel is achieved by input distributions whose coordinates are independent. To prove this, we will show that the rate constraints in (3.28) are maximized by *diagonal* matrices Λ_1 and Λ_2, and therefore, the matrices of optimum eigenvectors for Σ_1 and Σ_2 are V and U, respectively. First we apply the Hadamard inequality (the determinant of a nonnegative-definite matrix is upper-bounded by the product of its diagonal elements) to the individual rate constraints in (3.28):

$$\frac{1}{2}\log\det\left[I_n+\sigma^{-2}\Lambda_k\right]\le\frac{1}{2}\sum_{i=1}^{n}\log\left(1+\frac{\lambda_{ki}}{\sigma^2}\right),\qquad k=1,2 \quad (3.30)$$

where λ_{ki} is the ith diagonal entry of Λ_k, and equality holds in (3.30) when Λ_k is diagonal. Second, to upper-bound the rate-sum constraint in (3.28) in terms of the diagonal elements of Λ_1 and Λ_2, we will invoke the following result proved in Appendix III.

Lemma 2. Let A and B be $n\times n$ nonnegative-definite matrices, and let $\Delta=\operatorname{diag}\{\delta_1,\cdots,\delta_n\}$, where δ_i is a complex scalar such that $|\delta_i|\le 1$ for $i=1,\cdots,n$. Then

$$\det\left[I_{2n}+\begin{bmatrix}A & 0\\ 0 & B\end{bmatrix}\begin{bmatrix}I_n & \Delta^*\\ \Delta & I_n\end{bmatrix}\right]\le\prod_{i=1}^{n}\left\{1+a_{ii}+b_{ii}+a_{ii}b_{ii}(1-|\delta_i|)^2\right\} \quad (3.31)$$

with equality if A and B are diagonal.

We apply Lemma 2 to the case $A=\sigma^{-2}\Lambda_1$, $B=\sigma^{-2}\Lambda_2$ and $\Delta=D$, where the singular values of S (i.e., the diagonal elements of D) are real numbers belonging to the interval $(0,1)$ since R is positive-definite. (See [19, p. 382], and (II.7).) It then follows from (3.30) and Lemma 2 that the three constraints in (3.28) are maximized by diagonal matrices; hence we can now write

$$\frac{1}{n}C_n=\bigcup_{\substack{\lambda_{ki}\ge 0,\ i=1,\cdots n\\ \frac{1}{n}\sum_{i=1}^{n}\lambda_{ki}\le w_k\\ k=1,2}}\left\{0\le R_1\le\frac{1}{2n}\sum_{i=1}^{n}\log\left(1+\frac{\lambda_{1i}}{\sigma^2}\right)\right.$$

$$0\le R_2\le\frac{1}{2n}\sum_{i=1}^{n}\log\left(1+\frac{\lambda_{2i}}{\sigma^2}\right)$$

$$\left.R_1+R_2\le\frac{1}{2n}\sum_{i=1}^{n}\log\left(1+\frac{\lambda_{1i}}{\sigma^2}+\frac{\lambda_{2i}}{\sigma^2}+\frac{\lambda_{1i}\lambda_{2i}}{\sigma^4}\left(1-d_i^2\right)\right)\right\}. \quad (3.32)$$

It remains to show that the limit as $n\to\infty$ of the set (3.32) is equal to (3.13). The approach we follow is to show that the Pareto-optimal[5] rate pairs of (3.13) coincide with those of $\lim_{n\to\infty}(1/n)C_n$.

The integrand in the rate-sum constraint of the region C in (3.13) is equal to (cf.(3.9))

$$\log\det\left[I_2+\frac{1}{\sigma^2}\begin{bmatrix}S_1(\omega) & 0\\ 0 & S_2(\omega)\end{bmatrix}T(\omega)\right]$$

which is a concave function of $(S_1(\omega),S_2(\omega))$ for all $\omega\in[-\pi,\pi]$. Then C can be shown to be a convex set by following the reasoning we used to show the convexity of C_n. However, if the closed set C is convex, then each of its Pareto-optimal rate pairs has the property that it attains the maximum

$$\max_{(R_1,R_2)\in C}\alpha R_1+(1-\alpha)R_2 \quad (3.33)$$

for some $0\le\alpha\le 1$ (see [20]).

For each spectral pair, denote the rate-sum constraint in (3.13) by

$$F(S_1,S_2)=\frac{1}{4\pi}\int_{-\pi}^{\pi}\log\left(1+\frac{S_1(\omega)}{\sigma^2}+\frac{S_1(\omega)}{\sigma^2}+\frac{S_1(\omega)S_2(\omega)}{\sigma^4}\left(1-\rho^2(\omega)\right)\right)d\omega. \quad (3.34)$$

Notice that the individual rate constraints in (3.13) are $F(S_1,0)$ and $F(0,S_2)$, respectively. Furthermore, to simplify the notation, the $L_1[-\pi,\pi]$ subset of power spectra satisfying $(1/2\pi)\int_{-\pi}^{\pi}S(\omega)\,d\omega\le w_i$ will be denoted by

[5] An achievable rate pair (R_1,R_2) is Pareto-optimal if no other pair $(R_1+\delta_1,R_2+\delta_2)$ with $\delta_1\ge 0$ and $\delta_2>0$ is achievable. For example, in the pentagons (2.7) and (3.6) only the points on the boundary of the capacity region belonging to the 135° segment are Pareto optimal

$P(w)$. Then for every $0 \le \alpha \le 1$, (3.33) is equal to

$$\max_{\substack{S_1 \in P(w_1) \\ S_2 \in P(w_2)}} \max_{\substack{0 \le R_1 \le F(S_1,0) \\ 0 \le R_2 \le F(0,S_2) \\ R_1+R_2 \le F(S_1,S_2)}} \alpha R_1 + (1-\alpha) R_2$$

$$= \max_{\substack{S_1 \in P(w_1) \\ S_2 \in P(w_2)}} \max \{ \alpha F(S_1,0) + (1-\alpha)[F(S_1,S_2) - F(S_1,0)],\ \alpha[F(S_1,S_2) - F(0,S_2)] + (1-\alpha) F(0,S_2) \}$$

$$= \begin{cases} \max\limits_{\substack{S_1 \in P(w_1) \\ S_2 \in P(w_2)}} (2\alpha-1) F(S_1,0) + (1-\alpha) F(S_1,S_2), & \text{if } \frac{1}{2} \le \alpha \le 1 \\ \max\limits_{\substack{S_1 \in P(w_1) \\ S_2 \in P(w_2)}} (1-2\alpha) F(0,S_2) + \alpha F(S_1,S_2), & \text{if } 0 \le \alpha \le \frac{1}{2} \end{cases} \tag{3.35}$$

where (3.35) follows from

$$F(S_1,S_2) \le F(S_1,0) + F(0,S_2). \tag{3.36}$$

Following the same approach with the convex set (3.32), we obtain that every Pareto-optimal pair in $(1/n)C_n$ attains

$$\max_{(R_1,R_2) \in \frac{1}{n} C_n} \alpha R_1 + (1-\alpha) R_2$$

$$= \begin{cases} \max\limits_{\substack{\lambda_{kl} \ge 0, l=1,\cdots,n \\ \frac{1}{n}\sum_{l=1}^{n} \lambda_{kl} \le w_k \\ k=1,2}} \frac{1}{2n} \sum_{l=1}^{n} (2\alpha-1) \log\left(1 + \frac{\lambda_{1l}}{\sigma^2}\right) + (1-\alpha) \log\left(1 + \frac{\lambda_{1l}}{\sigma^2} + \frac{\lambda_{2l}}{\sigma^2} + \frac{\lambda_{1l}\lambda_{2l}}{\sigma^4}(1-d_l^2)\right), & \text{if } \frac{1}{2} \le \alpha \le 1 \\ \max\limits_{\substack{\lambda_{kl} \ge 0, l=1,\cdots,n \\ \frac{1}{n}\sum_{l=1}^{n} \lambda_{kl} \le w_k \\ k=1,2}} \frac{1}{2n} \sum_{l=1}^{n} (1-2\alpha) \log\left(1 + \frac{\lambda_{2l}}{\sigma^2}\right) + \alpha \log\left(1 + \frac{\lambda_{1l}}{\sigma^2} + \frac{\lambda_{2l}}{\sigma^2} + \frac{\lambda_{1l}\lambda_{2l}}{\sigma^4}(1-d_l^2)\right), & \text{if } 0 \le \alpha \le \frac{1}{2} \end{cases} \tag{3.37}$$

for some $0 \le \alpha \le 1$. To show that, for every $0 \le \alpha \le 1$, the limit as $n \to \infty$ of the right side of (3.37) coincides with (3.35), we will fix $1/2 \le \alpha \le 1$ (the proof for $0 \le \alpha < 1/2$ is identical) and we will prove that

$$\max_{\substack{\Phi_1 \in P\left(\frac{w_1}{\sigma^2}\right) \\ \Phi_2 \in P\left(\frac{w_2}{\sigma^2}\right)}} \frac{1}{4\pi} \int_{-\pi}^{\pi} f\left(\Phi_1(\omega), \Phi_2(\omega), \rho^2(\omega)\right) d\omega$$

$$= \lim_{n \to \infty} \max_{\substack{\phi_{kl} \ge 0, l=1,\cdots,n \\ \frac{1}{n}\sum_{l=1}^{n} \phi_{kl} \le \frac{w_k}{\sigma^2} \\ k=1,2}} \frac{1}{2n} \sum_{l=1}^{n} f\left(\phi_{1l}, \phi_{2l}, d_l^2\right), \tag{3.38}$$

where

$$f(z_1, z_2, \rho^2) = (2\alpha-1)\log(1+z_1) + (1-\alpha)\log\left(1 + z_1 + z_2 + z_1 z_2 (1-\rho^2)\right). \tag{3.39}$$

If $\alpha = 1$, then (3.38) follows immediately because the maximizing arguments in the left and right sides therein are easily shown to be the constant functions $\Phi_k(\omega) = \phi_{kl} = w_k/\sigma^2$, $\omega \in [-\pi, \pi]$, $i = 1, \cdots, n$, $k = 1,2$. If $1/2 \le \alpha < 1$, we invoke the following result (proved in Appendix IV) regarding the optimization problem in (3.38).

Lemma 3: If $1/2 \le \alpha < 1$, then

$$\max_{\substack{\Phi_1 \in P\left(\frac{w_1}{\sigma^2}\right) \\ \Phi_2 \in P\left(\frac{w_2}{\sigma^2}\right)}} \frac{1}{4\pi} \int_{-\pi}^{\pi} f\left(\Phi_1(\omega), \Phi_2(\omega), \rho^2(\omega)\right) d\omega = \frac{1}{4\pi} \int_{-\pi}^{\pi} g\left(\rho^2(\omega), \theta_1, \theta_2\right) d\omega \tag{3.40}$$

where θ_1, θ_2 are positive scalars such that

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \gamma_k\left(\rho^2(\omega), \theta_1, \theta_2\right) d\omega = \frac{w_k}{\sigma^2}, \qquad k = 1,2 \tag{3.41}$$

$$g(\rho^2, \theta_1, \theta_2) = f\left(\gamma_1(\rho^2, \theta_1, \theta_2), \gamma_2(\rho^2, \theta_1, \theta_2), \rho^2\right) \tag{3.42}$$

and $\gamma_k(\cdot, \theta_1, \theta_2)$, $k = 1,2$ are continuous functions.

Proceeding as in the proof of Lemma 3, it follows that the same result holds for the finite-dimensional optimiza-

tion problem in the right side of (3.38), i.e.,

$$\max_{\substack{\phi_{ki}\geq 0 i=1,\cdots,n \\ \frac{1}{n}\sum_{i=1}^{n}\phi_{ki}\leq \frac{w_k}{\sigma^2} \\ k=1,2}} \frac{1}{2n}\sum_{i=1}^{n} f(\phi_{1i},\phi_{2i},d_i^2) = \frac{1}{2n}\sum_{i=1}^{n} g(d_i^2,\theta_1,\theta_2) \tag{3.43}$$

with θ_1, θ_2 such that

$$\frac{1}{n}\sum_{i=1}^{n}\gamma_k(d_i^2,\theta_1,\theta_2) = \frac{w_k}{\sigma^2}, \qquad k=1,2. \tag{3.44}$$

Since for any pair of signal-to-noise ratios $(w_1/\sigma^2, w_2/\sigma^2)$ there exist solutions θ_1 and θ_2 to (3.41) and to (3.44), the identity in (3.38) will follow if we can show that for every fixed positive pair (θ_1,θ_2)

$$\lim_{n\to\infty}\frac{1}{2n}\sum_{i=1}^{n} g(d_i^2,\theta_1,\theta_2) = \frac{1}{4\pi}\int_{-\pi}^{\pi} g(\rho^2(\omega),\theta_1,\theta_2)\,d\omega \tag{3.45}$$

and

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\gamma_k(d_i^2,\theta_1,\theta_2) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\gamma_k(\rho^2(\omega),\theta_1,\theta_2)\,d\omega. \tag{3.46}$$

To prove this, we need to examine the behavior as $n\to\infty$ of $(d_1^2,\cdots,d_n^2)$, the eigenvalues of the Jacobi matrix S^TS in (3.25). It can be shown that

$$d_i^2 = \rho_{12}^2+\rho_{21}^2+2\rho_{12}\rho_{21}\beta_i, \qquad i=1,\cdots,n \tag{3.47}$$

where $\{\beta_i,\ i=1,\cdots,n\}$ are the roots of the nth degree polynomial $T_{n+1}(x)$ obtained through the recursion[6]

$$T_{k+1}(x) = 2xT_k(x)-T_{k-1}(x)$$
$$T_1(x)=1; \quad T_0(x) = -\frac{\rho_{21}}{\rho_{12}}.$$

In special cases it is indeed possible to obtain closed-form expressions for the eigenvalues of S^TS, for example, if $\rho_{12}=\rho_{21}$, then [22]

$$d_i^2 = \rho_{12}^2+\rho_{21}^2+2\rho_{12}\rho_{21}\cos\left(\frac{2\pi i}{2n+1}\right), \qquad i=1,\cdots,n. \tag{3.48}$$

At any rate, it is easy to show that the eigenvalues of the Toeplitz matrix T obtained by substituting the entry $(S^TS)_{11}=\rho_{12}^2$ by $\rho_{12}^2+\rho_{21}^2$ are equal to [22]

$$\hat{d}_i^2 = \rho_{21}^2+\rho_{12}^2+2\rho_{12}\rho_{21}\cos\left(\frac{i\pi}{n+1}\right), \qquad i=1,\cdots,n. \tag{3.49}$$

Thus if d_i^2 were replaced by $\hat{d}_i^2$ in the left sides of (3.45) and (3.46), these equations would follow immediately because of the continuity of the Riemann integrands therein. It is indeed valid to carry out this replacement because T and S^TS differ in only one entry and are uniformly bounded in operator norm; thus they are asymptotically (as their dimension grows) equivalent [23], i.e., S^TS is asymptotically Toeplitz, and since $g(\cdot,\theta_1,\theta_2)$ and $\gamma_k(\cdot,\theta_1,\theta_2)$ are continuous functions, their averages evaluated at $\{d_i^2\}_{i=1}^n$ and $\{\hat{d}_i^2\}_{i=1}^n$ coincide as $n\to\infty$ [23, Theorem 2.3]. Hence (3.45) and (3.46) hold and the proof of (3.38) is complete.

[6]Except for the initial conditions, this is the recursive definition of the Chebyshev polynomials, whose zeros are trigonometric [21].

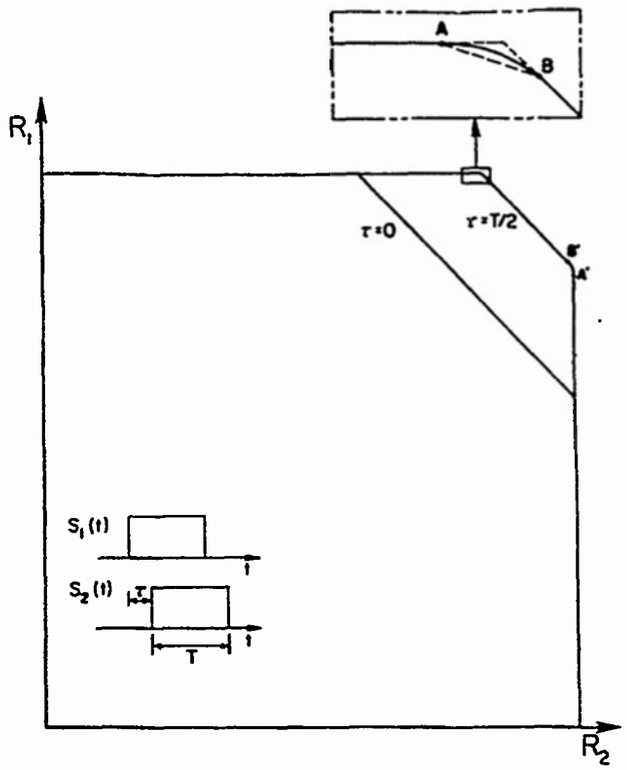

Fig. 5. Symbol-asynchronous Gaussian capacity region for rectangular signals and identical signal-to-noise ratios, when transmitters know offset between their signals.

Finally, note that Theorem 1 was proved under the assumption that the transmitters are frame synchronous. However, it follows from the results in [7] that the same capacity region holds even if the transmitters are frame-asynchronous because the capacity region is achieved by stationary distributions.

Fig. 5 shows the capacity region of the simplest possible symbol-asynchronous Gaussian multiple-access channel: the transmitters are assigned the same rectangular waveform and know the offset between their signals. The numerical computation of the capacity region is carried out using the results (IV.5–IV.13) of the functional optimization problem solved in the proof of Lemma 3. The worst-case offset between the signals is zero—in which case the channel is symbol synchronous and admits the scalar discrete-time model (2.6) resulting in the Cover–Wyner capacity region (2.7). The most favorable case occurs when the symbol offset is equal to half of the symbol period, in which case the outer region in Fig. 5 is the capacity region C in (3.13) computed with $\rho_{12}=\rho_{21}=0.5$. This capacity region, which is representative of that of any strictly asynchronous channel (i.e., when ρ_{12} and ρ_{21} are both nonzero),

resembles a pentagon with *smooth* corners. As we saw, the reason the region C is not a pentagon is that there is no unique pair of spectral densities in (3.13) that maximizes all rate constraints simultaneously. Consider the pentagon defined by points B and B' in Fig. 5. This is the subset of (3.13)

$$\{(R_1, R_2): 0 \le R_1 \le F(S_1^*,0), 0 \le R_2 \le F(0, S_2^*), \\ R_1 + R_2 \le F(S_1^*, S_2^*)\}$$

achievable with the unique spectral pair (S_1^*, S_2^*) that maximizes the rate-sum constraint, i.e.,

$$F(S_1^*, S_2^*) = \max_{\substack{S_1 \in P(w_1) \\ S_2 \in P(w_2)}} F(S_1, S_2) \tag{3.50}$$

and the rate-pairs B and B' correspond to $(R_1, R_2) = (F(S_1^*, 0), F(S_1^*, S_2^*) - F(S_1^*, 0))$ and $(R_1, R_2) = (F(S_1^*, S_2^*) - F(0, S_2^*), F(0, S_2^*))$, respectively.

Note that according to the optimality conditions in (IV.3), (IV.4) (particularized to $\alpha = 1/2$), the spectral pair (S_1^*, S_2^*) is the solution to

$$\frac{S_k^*(\omega)}{\sigma^2} = \Phi_k^*(\omega) \\ = \max\left\{\frac{1}{2\theta_k} - \frac{1+\Phi_j^*(\omega)}{1+\Phi_j^*(\omega)(1-\rho^2(\omega))}, 0\right\}, \\ (j,k) = (1,2),(2,1)$$

where (θ_1, θ_2) is chosen so that (IV.2) is satisfied. Since ρ_{12} and ρ_{21} are nonzero, $\rho^2(\omega)$ is not a constant function of ω, and hence neither are $S_1^*(\omega)$ and $S_2^*(\omega)$. However, the individual rate constraints

$$\frac{1}{4\pi}\int_{-\pi}^{\pi} \log\left(1 + \frac{S_k(\omega)}{\sigma^2}\right) d\omega$$

are maximized over the set $P(w_k)$ by the constant functions $S_k(\omega) = w_k$ and thus (S_1^*, S_2^*) fails to achieve the largest possible individual rates:

$$C_k = \frac{1}{2}\log\left(1 + \frac{w_k}{\sigma^2}\right), \qquad k = 1,2.$$

These rates are achieved (by one user at a time) at the points A and A' in Fig. 5. Point A is achieved by the spectral pair (w_1, S_2^a), where

$$F(w_1, S_2^a) = \max_{S_2 \in P(w_2)} F(w_1, S_2) \\ = \max_{S_2 \in P(w_2)} \frac{1}{4\pi}\int_{-\pi}^{\pi} \log\left(1 + \frac{w_1}{\sigma^2} + \frac{S_2(\omega)}{\sigma^2} \right. \\ \left. \cdot \left[1 + \frac{w_1}{\sigma^2}(1-\rho^2(\omega))\right]\right) d\omega \tag{3.51}$$

i.e., S_2^a is the best spectrum for user 2 when user 1 transmits at full single-user speed $(1/2)\log(1+(w_1/\sigma^2))$. The solution to (3.51) is

$$S_2^a(\omega) = \sigma^2 \max\left\{\beta - \frac{1+\frac{w_1}{\sigma^2}}{1+\frac{w_1}{\sigma^2}(1-\rho^2(\omega))}, 0\right\} \tag{3.52}$$

where β is chosen so that $1/2\pi \int S_2^a(\omega)\, d\omega = w_2$. Note that (3.52) admits the classical *water-filling* interpretation [24], [25] arising in the study of colored Gaussian single-user channel capacity.

The segment uniting A and B does not belong to the boundary of the capacity region, and therefore, C is not a heptagon. This property which is illustrated by the capacity region in Fig. 5 can be proved as follows. Choose $1/2 < \alpha^* < 1$ such that the rate pairs A and B (and their convex combinations) achieve the same value of the function $\alpha^* R_1 + (1-\alpha^*)R_2$, i.e. (cf. (3.35)),

$$(2\alpha^* - 1)F(w_1, 0) + (1-\alpha^*)F(w_1, S_2^a) \\ = (2\alpha^* - 1)F(S_1^*, 0) + (1-\alpha^*)F(S_1^*, S_2^*).$$

If the segment between A and B belonged to the boundary of the capacity region, then both A and B would attain $\max_{(R_1,R_2)\in C} \alpha^* R_1 + (1-\alpha^*)R_2$. However, this is not possible due to the strict concavity of the function $(2\alpha - 1)F(S_1, 0) + (1-\alpha)F(S_1, S_2)$: any convex combination of the spectral pairs (w_1, S_2^a) and (S_1^*, S_2^*) will achieve strictly higher values of $\alpha^* R_1 + (1-\alpha^*)R_2$ than A and B. In fact, the same argument can be employed to show that the transition from A to B contains no straight lines.

We are now ready to state and prove our main result concerning the capacity region of the asynchronous Gaussian multiple-access channel wherein the transmitters ignore their mutual offset $\tau_2 - \tau_1$. The transmitters only know that the crosscorrelations (ρ_{12}, ρ_{21}) that parametrize the channel belong to an uncertainty set Γ, which is determined by the choice of the signature waveforms. For example, if both users are assigned a rectangular waveform then the uncertainty set is equal to the segment $\Gamma = \{0 \le \rho_{12} \le 1, 0 \le \rho_{21} \le 1, \rho_{12} + \rho_{21} = 1\}$. Note that in practical applications it may be of interest to model channels where the offset is not the only source of uncertainty for the crosscorrelations; for example, if the signature waveforms are sinusoidally modulated, the crosscorrelations depend on the relative phase between the carriers (e.g., for rectangular signals modulated by sinusoids whose frequency is a large multiple of the inverse of the symbol period T, we get $\Gamma = \pm\{0 \le \rho_{12} \le 1, 0 \le \rho_{21} \le 1, \rho_{12} + \rho_{21} \le 1\}$). The following result puts no restrictions on the source of the uncertainty of the set of cross correlations Γ.

Theorem 2: The capacity region of the energy-constrained asynchronous Gaussian multiple-access channel where the transmitters do not know their mutual offset is

given by

$$C^* = \bigcup_{\substack{S_k(\omega)\ge 0,\ \omega\in[-\pi,\pi] \\ \frac{1}{2\pi}\int_{-\pi}^{\pi} S_k(\omega)\,d\omega = w_k \\ k=1,2}} \left\{ (R_1, R_2),\ 0 \le R_1 \le \frac{1}{4\pi}\int_{-\pi}^{\pi} \log\left(1+\frac{S_1(\omega)}{\sigma^2}\right) d\omega,\ 0 \le R_2 \le \frac{1}{4\pi}\int_{-\pi}^{\pi} \log\left(1+\frac{S_2(\omega)}{\sigma^2}\right) d\omega \right.$$

$$\left. R_1 + R_2 \le \inf_{(\rho_{12},\rho_{21})\in\Gamma} \frac{1}{4\pi}\int_{-\pi}^{\pi} \log\left(1+\frac{S_1(\omega)}{\sigma^2}+\frac{S_2(\omega)}{\sigma^2}+\frac{S_1(\omega)S_2(\omega)}{\sigma^4}\left[1-\rho_{12}^2-\rho_{21}^2-2\rho_{12}\rho_{21}\cos\omega\right]\right) d\omega \right\}. \quad (3.53)$$

Proof: Having shown the result for the special case of a singleton uncertainty set $\Gamma = \{(\rho_{12}, \rho_{21})\}$, we will be able to proceed at a faster tempo by invoking several lemmas used in the proof of Theorem 1. The capacity of the compound decoder-informed multiple-access channel with memory can be shown to satisfy ([7], see also [25, p. 288])

$$C^* = \text{closure}\left(\liminf_{n\to\infty} \frac{1}{n} C_n^* \right) \quad (3.54)$$

with

$$C_n^* = \bigcup_{X_1^n,\, X_2^n} \bigcap_{(\rho_{12},\rho_{21})\in\Gamma} \left\{ \begin{array}{l} (R_1^n, R_2^n): \\ \\ \\ \end{array} \begin{array}{l} 0 \le R_1^n \le I(X_1^n; Y^n(\rho_{12},\rho_{21}) | X_2^n) \\ 0 \le R_2^n \le I(X_2^n; Y^n(\rho_{12},\rho_{21}) | X_1^n) \\ R_1^n + R_2^n \le I(X_1^n, X_2^n; Y^n(\rho_{12},\rho_{21})) \end{array} \right\} \quad (3.55)$$

where $Y^n(\rho_{12}, \rho_{21})$ denotes the output of channel (2.4) with crosscorrelations (ρ_{12}, ρ_{21}), and X_1^n and X_2^n are independent random variables satisfying the same input constraints as in (3.15). This follows simply because the direct coding theorem can be proved using codebooks that do not depend on the actual channel (via random selection) while the fact that for reliable communication a code has to be good no matter which actual channel is in effect establishes the converse theorem. Using Lemma 1 and proceeding as in Theorem 1, we obtain that

$$C_n^* = \bigcup_{\substack{\Sigma_k \ge 0 \\ \frac{1}{n}\text{tr}\,\Sigma_k \le w_k \\ k=1,2}} \left\{ (R_1, R_2),\ 0 \le R_1 \le \frac{1}{2}\log\det\left[I_n + \sigma^{-2}\Sigma_1\right] \right.$$

$$\left. 0 \le R_2 \le \frac{1}{2}\log\det\left[I_n + \sigma^{-2}\Sigma_2\right] \right.$$

$$\left. R_1 + R_2 \le \inf_{(\rho_{12},\rho_{21})\in\Gamma} \frac{1}{2}\log\det\left[I_{2n} + \frac{1}{\sigma^2}\begin{bmatrix}\Sigma_1 & 0 \\ 0 & \Sigma_2\end{bmatrix}\begin{bmatrix} I_n & S^T \\ S & I_n\end{bmatrix}\right] \right\} \quad (3.56)$$

where the only difference with respect to (3.23) is the minimization of the rate-sum constraint with respect to the cross-correlation-dependent matrix S.

In Theorem 1, we showed using the singular-value decomposition theorem that a set of eigenvectors exists that maximizes the three constraints in (3.23) no matter which eigenvalues are used, thus reducing the union therein to one over diagonal matrices. Here this property is no longer true, and we need to follow an alternative route, suggested by the following result.

Lemma 4: Define the circulant matrix

$$\tilde{S} = \rho_{12} I_n + \rho_{21}\begin{bmatrix} 0 & 1 & 0 & & & \\ 0 & 0 & 1 & & & \\ & 0 & 0 & \vdots & 0 & \\ & & & & 1 & 0 \\ & & & & 0 & 1 \\ 1 & & & & 0 & 0 \end{bmatrix}$$

which differs from S only in the $(n,1)$ entry (cf. (3.22)). Then for every $\delta > 0$, and $n > n_\delta$ (independent of Σ_1, Σ_2, ρ_{12} and ρ_{21})

$$\left| \frac{1}{n}\log\det\left[I_{2n} + \frac{1}{\sigma^2}\begin{bmatrix}\Sigma_1 & 0 \\ 0 & \Sigma_2\end{bmatrix}\begin{bmatrix} I_n & \tilde{S}^T \\ \tilde{S} & I_n\end{bmatrix}\right] \right.$$

$$\left. - \frac{1}{n}\log\det\left[I_{2n} + \frac{1}{\sigma^2}\begin{bmatrix}\Sigma_1 & 0 \\ 0 & \Sigma_2\end{bmatrix}\begin{bmatrix} I_n & S^T \\ S & I_n\end{bmatrix}\right] \right| < \delta. \quad (3.57)$$

Therefore, as far as computing $\lim_{n\to\infty}(1/n)C_n^*$ is concerned, we can substitute S by $\tilde{S}$ in (3.56). The effect of this substitution is to introduce an artificial interference term between symbol 1 of user 1 and symbol n of user 2 (Fig. 6), resulting in a channel which can be thought of as a wrapped-around version of channel (2.1). By the circular symmetry of this new channel, we can intuitively expect the covariances achieving capacity to be circulant, and consequently, the existence of a set of optimum eigenvectors (whose components are powers of the complex roots of untity) which do not depend on the crosscorrelations. To show this, it suffices to write

$$\tilde{S} = \tilde{U}\tilde{D}\tilde{U}^* \quad (3.58)$$

where $\tilde{D}$ is a diagonal matrix of the eigenvalues of $\tilde{S}$, which coincide with the DFT of the first row of the circulant matrix $\tilde{S}$ [23]

$$\tilde{d}_k = \rho_{12} + \rho_{21} e^{-j2\pi(k-1)/n}, \qquad k = 1, \cdots, n \quad (3.59)$$

and $\tilde{U}$ is the orthonormal matrix of eigenvectors of $\tilde{S}$ given by

$$\tilde{U}_{ik} = \frac{1}{\sqrt{n}} e^{-j2\pi(k-1)(i-1)/n}, \qquad i = 1, \cdots, n,\ k = 1, \cdots, n. \quad (3.60)$$

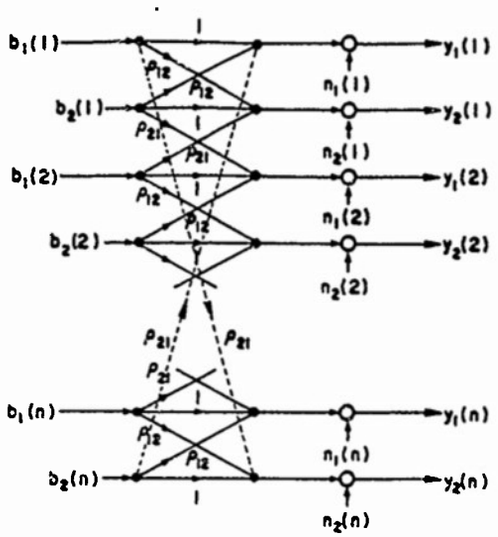

Fig. 6. Circular n-block memoryless two-user channel.

Using the decomposition (3.58) in lieu of (3.24) and Lemma 2 with $\Delta = \tilde{D}$, and proceeding in a way similar to the proof of Theorem 1, we obtain

$$\bigcup_{\substack{\Sigma_k \ge 0 \\ \frac{1}{n}\mathrm{tr}\Sigma_k \le w_k \\ k=1,2}} \left\{ (R_1, R_2), 0 \le R_1 \le \frac{1}{2n}\log\det\left[I_n + \sigma^{-2}\Sigma_1\right], 0 \le R_2 \le \frac{1}{2n}\log\det\left[I_n + \sigma^{-2}\Sigma_2\right] \right.$$

$$\left. R_1 + R_2 \le \inf_{(\rho_{12},\rho_{21})\in\Gamma} \frac{1}{2n}\log\det\left[I_{2n} + \frac{1}{\sigma^2}\begin{bmatrix}\Sigma_1 & 0\\ 0 & \Sigma_2\end{bmatrix}\begin{bmatrix}I_n & S^T\\ S & I_n\end{bmatrix}\right] \right\}$$

$$= \bigcup_{\substack{\lambda_{ki}\ge 0, i=1,\cdots,n \\ \frac{1}{n}\sum_{i=1}^{n}\lambda_{ki} \le w_k \\ k=1,2}} \left\{ (R_1, R_2), 0 \le R_1 \le \frac{1}{2n}\sum_{i=1}^{n}\log\left(1+\frac{\lambda_{1i}}{\sigma^2}\right),\ 0 \le R_2 \le \frac{1}{2n}\sum_{i=1}^{n}\log\left(1+\frac{\lambda_{2i}}{\sigma^2}\right) \right.$$

$$\left. R_1 + R_2 \le \inf_{(\rho_{12},\rho_{21})\in\Gamma} \frac{1}{2n}\sum_{i=1}^{n}\log\left(1+\frac{\lambda_{1i}}{\sigma^2}+\frac{\lambda_{2i}}{\sigma^2}+\frac{\lambda_{1i}\lambda_{2i}}{\sigma^4}\left(1-\left|\tilde{d}_i\right|^2\right)\right) \right\} \quad (3.61)$$

where

$$\left|\tilde{d}_i\right|^2 = \rho_{12}^2 + \rho_{21}^2 + 2\rho_{12}\rho_{21}\cos\frac{2\pi(i-1)}{n} = \rho^2\left(\frac{2\pi(i-1)}{n}\right). \quad (3.62)$$

As we saw in the proof of Theorem 1, the convergence of the right side of (3.61) to (3.53) reduces to

$$\max_{\Phi_1 \in P_n\left(\frac{w_1}{\sigma^2}\right)} \inf_{(\rho_{12},\rho_{21})\in\Gamma} \frac{1}{4\pi}\int_{-\pi}^{\pi} f\left(\Phi_1(\omega), \Phi_2(\omega), \rho^2(\omega)\right) d\omega$$

$$= \lim_{n\to\infty} \max_{\substack{\phi_{ki}\ge 0, i=1,\cdots,n \\ \frac{1}{n}\sum_{i=1}^{n}\phi_{ki} \le \frac{w_k}{\sigma^2} \\ k=1,2}} \inf_{(\rho_{12},\rho_{21})\in\Gamma} \frac{1}{2n}\sum_{i=1}^{n} f\left(\phi_{1i}, \phi_{2i}, \rho^2\left(\frac{2\pi(i-1)}{n}\right)\right), \quad (3.63)$$

However, note that the right side of (3.63) can be written as

$$\lim_{n\to\infty} \max_{\substack{\Phi_1 \in P_n\left(\frac{w_1}{\sigma^2}\right) \\ \Phi_2 \in P_n\left(\frac{w_2}{\sigma^2}\right)}} \inf_{(\rho_{12},\rho_{21})\in\Gamma} \frac{1}{4\pi} \cdot \int_{-\pi}^{\pi} f\left(\Phi_1(\omega), \Phi_2(\omega), \rho_n^2(\omega)\right) d\omega \quad (3.64)$$

where

$$\rho_n^2(\omega) = \rho^2\left(\frac{2\pi(i-1)}{n}\right), \qquad \text{for } \omega \in \left[\frac{2\pi(i-1)}{n}, \frac{2\pi i}{n}\right]$$

and $P_n(w_k/\sigma^2)$ is the subset of $P(w_k/\sigma^2)$ of piecewise constant functions on the partition $[0, 2\pi/n, \cdots, 2\pi]$. Since $\rho_n^2(\omega)$ is piecewise constant on that partition, it is easy to show that we can replace $P_n(w_k/\sigma^2)$ by $P(w_k/\sigma^2)$ in (3.64) without changing the maximum value for any n. Finally, (3.63) follows from the fact that for every $\epsilon > 0$, and $n > n_\epsilon$ (independent of Φ_1, Φ_2, ρ_{12}, and ρ_{21})

$$\left| \int_{-\pi}^{\pi} f\left(\Phi_1(\omega), \Phi_2(\omega), \rho_n^2(\omega)\right) d\omega - \int_{-\pi}^{\pi} f\left(\Phi_1(\omega), \Phi_2(\omega), \rho^2(\omega)\right) d\omega \right| \le 2\pi\log\left(1+\frac{w_1 w_2}{\sigma^4}\epsilon\right). \quad (3.65)$$

Again notice that since the capacity region is achieved with stationary inputs, Theorem 2 holds regardless of whether or not the transmitters are frame synchronous.

Corollary: If both users are assigned identical waveforms (and they do not know their mutual offset), then the capacity region is invariant to symbol (and frame) asynchronism.

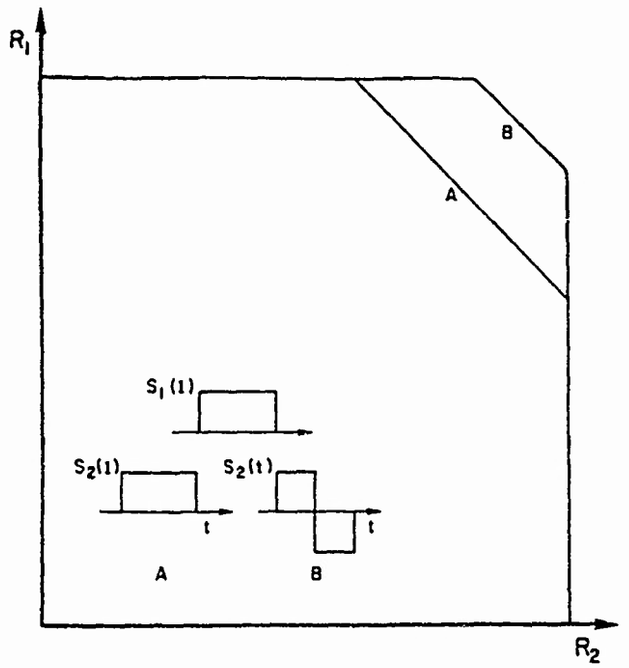

Fig. 7. Symbol-asynchronous Gaussian capacity region when transmitters do not know offset between their signals.

Proof: Because $1-\rho_{12}^2-\rho_{21}^2-2\rho_{12}\rho_{21}\cos\omega \geq 0$, it is easy to see that the asynchronous capacity region (3.53) reduces to the Cover–Wyner region (2.7) if the uncertainty set Γ includes either $(0,1)$ or $(1,0)$. This occurs when both users are assigned the same waveform.

In Fig. 7 we can see the asynchronous capacity regions corresponding to two different assignments of the signature waveform: a) identical signals, resulting in the Cover–Wyner pentagon, and b) signals that are orthogonal when symbol synchronous, resulting in a pentagon with *smooth* corners.

IV. Efficiency Region

A fruitful way to represent the multiple-access capacity region is to consider the *effective* signal-to-noise ratio of a user who transmits at rate R, which we define as the signal-to-noise ratio, γ, required to achieve capacity R in a single-user channel, i.e.,

$$\gamma = \exp[2R]-1. \tag{4.1}$$

Since the mapping in (4.1) is one-to-one, the rate and the effective signal-to-noise ratio give the same information. It is convenient to normalize the effective signal-to-noise ratio with respect to the actual signal-to-noise ratio. This results in the performance measure we refer to as efficiency[7] η, which is a parameter ranging from 0 to 1 that quantifies the performance degradation suffered by each user because of the presence of other users in the channel. Once the capacity region of a multiple-access channel is known it is immediate to obtain the efficiency

[7] An analogous performance measure was defined in the analysis of the minimum uncoded error probability of Gaussian multiple-access channels [27].

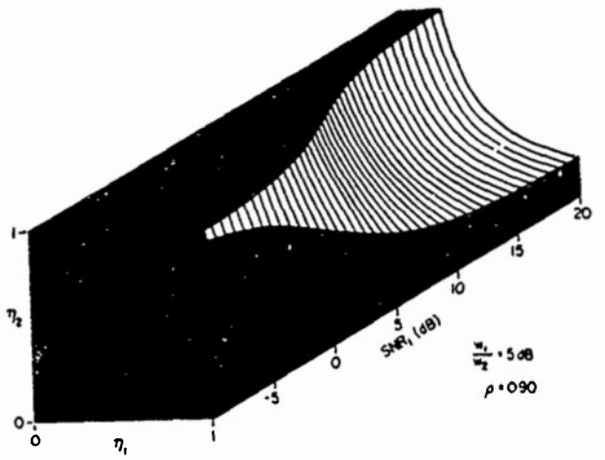

Fig. 8. Efficiency region as function of background noise level.

region, by substituting each of the individual rates in terms of the respective efficiencies, i.e.,

$$R_i = \frac{1}{2}\log\left(1+\eta_i\frac{w_i}{\sigma^2}\right). \tag{4.2}$$

For example, it follows from the capacity region in (3.6) that the efficiency region of the symbol-synchronous channel is equal to

$$\left\{(\eta_1,\eta_2): 0\leq\eta_1\leq 1, 0\leq\eta_2\leq 1, \right.$$
$$\left.\eta_1\eta_2-(1-\eta_1)\frac{\sigma^2}{w_2}-(1-\eta_2)\frac{\sigma^2}{w_1}\leq 1-\rho^2\right\} \tag{4.3}$$

where recall that ρ is the crosscorrelation between the assigned waveforms.

This efficiency region is illustrated in Fig. 8, as a function of the background noise level. For low signal-to-noise ratios the efficiency region occupies nearly all of the unit square because the main mechanism limiting performance is the background Gaussian noise, rather than the multiple-access interference. Conversely, it is apparent from Fig. 8 that for moderate-to-large signal-to-noise ratios the efficiency region converges to an asymptotic region which quantifies the underlying limitation of the multiple-access channel due to the cross correlation between the assigned signal waveforms. The region in (4.3) admits a particularly simple asymptotic expression as the noise spectral density goes to zero:

$$E = \left\{(\eta_1,\eta_2): 0\leq\eta_1\leq 1, 0\leq\eta_2\leq 1,\ \eta_1\eta_2\leq 1-\rho^2\right\}. \tag{4.4}$$

The usefulness of the asymptotic efficiency region is threefold: it provides a simple way to characterize multiple-access capacity in high signal-to-noise ratio situations; it gives a lower bound[8] to the efficiency region achievable at any background noise level, which depends only on the assigned signal waveforms and not on the signal-to-noise ratios, and it gives an intuitive characterization of the performance degradation in a multiuser channel in terms

[8] It is shown in [14] that the efficiency region is monotonically increasing in σ^2 (cf. Fig. 8).

Letting $\sigma^2 \to 0$, the asymptotic efficiency region results:

$$E = \bigcup_{\substack{S_i(\omega)\ge 0 \\ \frac{1}{2\pi}\int S_i(\omega)\,d\omega = 1 \\ i=1,2}} \left\{(\eta_1,\eta_2):\ \eta_1 \le \exp\frac{1}{2\pi}\int_{-\pi}^{\pi}\log S_1(\omega)\,d\omega,\ \eta_2 \le \exp\frac{1}{2\pi}\int_{-\pi}^{\pi}\log S_2(\omega)\,d\omega\right.$$

$$\left.\eta_1\eta_2 \le \exp\frac{1}{2\pi}\int_{-\pi}^{\pi}\log S_1(\omega)\,d\omega \exp\frac{1}{2\pi}\int_{-\pi}^{\pi}\log S_2(\omega)\,d\omega \inf_{(\rho_{12},\rho_{21})\in\Gamma}\exp\frac{1}{2\pi}\int_{-\pi}^{\pi}\log\left(1-\rho^2(\omega)\right)d\omega\right\} \quad (4.6)$$

of the additional power required to achieve single-user capacity.

For example, suppose that the users' objective is to transmit at rates $\bar{R}_1$ and $\bar{R}_2$, respectively. If they were operating in a single-user channel, these rates could be achieved with powers: $w_k = \sigma^2(\exp[2\bar{R}_k]-1)$, $k=1,2$. However, when they share the same channel, these powers are no longer sufficient to guarantee reliable communication at rates $\bar{R}_1$ and $\bar{R}_2$. The asymptotic efficiency region (4.4) indicates that the sum of their powers in dB has to increase by $-10\log(1-\rho^2)$ dB and that the way the users split the burden of increasing their powers is immaterial as long as the total power increases by the prescribed amount.

In the conventional scalar multiple-access channel, which corresponds to the users being assigned the same waveforms, i.e., $\rho = 1$, the asymptotic efficiency region is (via (4.4)):

$$\{0 \le \eta_1 \le 1,\ \eta_2 = 0\} \cup \{\eta_1 = 0,\ 0 \le \eta_2 \le 1\}.$$

Thus when the signal-to-noise ratio is high, the best strategy is to let one of the users transmit at practically full single-user speed, while the other user's rate is kept at a very low level. This is considerably more efficient than time-division multiple-access (TDMA) signaling whose asymptotic efficiency is equal to zero for both users—although if both rates are required to be the same, then TDMA is indeed almost as good as the best coding for low background noise (see [28]).

These conclusions do not hold in the case where the assigned waveforms are different ($|\rho| < 1$). For example, suppose that $\rho = 0.1$ and two equal-rate equal-energy users with signal-to-noise ratio equal to 20 dB transmit at the maximum possible rates. Had the users employed TDMA, each of them would have required approximately 40 dB to attain the same rate. Even in the case where there is heavy cross correlation between the signals, TDMA is not near-optimum, e.g., if $\rho = 0.9$, then TDMA would still require 33 dB to attain the same rate.

The efficiency region of the asynchronous Gaussian multiple-access channel is (via (3.53) and (4.2)) equal to

$$\bigcup_{\substack{S_i(\omega)\ge 0 \\ \frac{1}{2\pi}\int S_i(\omega)\,d\omega = 1 \\ i=1,2}} \left\{(\eta_1,\eta_2):\ \left(\eta_1+\frac{\sigma^2}{w_1}\right) \le \exp\frac{1}{2\pi}\int_{-\pi}^{\pi}\log\left[\frac{\sigma^2}{w_1}+S_1(\omega)\right]d\omega,\ \left(\eta_2+\frac{\sigma^2}{w_2}\right) \le \exp\frac{1}{2\pi}\int_{-\pi}^{\pi}\log\left[\frac{\sigma^2}{w_2}+S_2(\omega)\right]d\omega\right.$$

$$\left.\left(\eta_1+\frac{\sigma^2}{w_1}\right)\left(\eta_2+\frac{\sigma^2}{w_2}\right) \le \inf_{(\rho_{12},\rho_{21})\in\Gamma}\exp\frac{1}{2\pi}\int_{-\pi}^{\pi}\log\left[\left(S_1(\omega)+\frac{\sigma^2}{w_1}\right)\left(S_2(\omega)+\frac{\sigma^2}{w_2}\right)-\rho^2(\omega)S_1(\omega)S_2(\omega)\right]d\omega\right\} \quad (4.5)$$

The constraints in (4.6) depend on the spectral densities only through their geometric mean; therefore, all three constraints are maximized simultaneously by a single pair of spectral densities because the function that maximizes the geometric mean subject to a constraint on its arithmetic mean is constant. Therefore, (4.6) is equal to the efficiency region achievable with white spectral densities $S_k(\omega) = 1$, $\omega \in [-\pi, \pi]$, which implies that white inputs, while not optimum in general, achieve capacity asymptotically as the background noise level goes to zero. Then the asymptotic efficiency region is the intersection of the unit square with the hyperbolic region

$$\eta_1\eta_2 \le \inf_{(\rho_{12},\rho_{21})\in\Gamma}\exp\frac{1}{2\pi}\cdot\int_{-\pi}^{\pi}\log\left[1-\left(\rho_{12}^2+\rho_{21}^2\right)-2\rho_{12}\rho_{21}\cos\omega\right]d\omega$$

$$= \inf_{(\rho_{12},\rho_{21})\in\Gamma}\left[\frac{\sqrt{1-(\rho_{12}+\rho_{21})^2}+\sqrt{1-(\rho_{12}-\rho_{21})^2}}{2}\right]^2 \quad (4.7)$$

where the definite integral is found in [29, p. 560] (see also [30, p. 384]). Note that this result generalizes the constraint $1-\rho^2$ obtained for the product of the asymptotic efficiencies in (4.4) when the users are synchronous. Equation (4.7) indicates that contrary to what is sometimes assumed in pseudonoise sequence design, it is as important to minimize the *difference* between the cross correlations as to minimize their *sum* (the so-called periodic cross correlation). The function on the right side of (4.7) is tightly approximated by $1-\rho_{12}^2-\rho_{21}^2$ for low cross-correlation values such as those in Fig. 9, where we can see the uncertainty set of cross correlations between two carrier-modulated spread spectrum waveforms used in CDMA [31]. In this case, the minimization in (4.7) is attained by the rightmost point in Fig. 9.

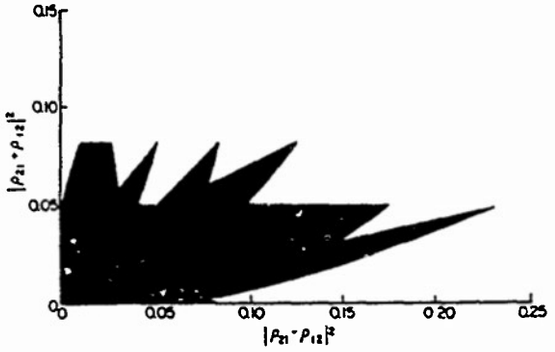

Fig. 9. Locus of cross correlations for maximal-length pseudonoise sequences with 31 chips per symbol.

APPENDIX I

To simplify the coding theorems for multiple-access channels with memory invoked to find the capacity region of the asynchronous Gaussian channel, it is convenient to obtain a discrete-time channel equivalent to (2.4) and whose noise process is independent. The idea is to obtain a set of sufficient statistics that are independent given the transmitted symbols, but that unlike those in (2.4) are not a minimal set. We define (cf. Fig. 2):

$$y_2^R(i) = \int_{(i+1)T+\tau_1}^{(i+1)T+\tau_2} y(t)s_2(t-iT-\tau_2)\,dt \qquad \text{(I.1a)}$$

$$y_2^L(i) = \int_{iT+\tau_2}^{(i+1)T+\tau_1} y(t)s_2(t-iT-\tau_2)\,dt. \qquad \text{(I.1b)}$$

It is clear from (2.2) and (I.1) that

$$y_2(i) = y_2^R(i) + y_2^L(i).$$

Thus the set of quantities $\{y_1(i)\}_{i=1}^n$, $\{y_2^R(i)\}_{i=1}^n$, and $\{y_2^L(i)\}_{i=1}^n$ are sufficient statistics for the transmitted messages. To obtain the explicit dependence of $y_2^R(i)$ and $y_2^L(i)$ on the transmitted symbols, it is convenient to define the partial energies of $\{s_2(t)\}$:

$$e_2^R = \int_{\tau+\tau_1-\tau_2}^{T} s_2^2(t)\,dt \qquad \text{(I.2a)}$$

$$e_2^L = \int_0^{T+\tau_1-\tau_2} s_2^2(t)\,dt. \qquad \text{(I.2b)}$$

Thus $e_2^R + e_2^L = 1$, and it follows from (2.1), (2.4), (I.1), and (I.2) that we can write

$$\begin{bmatrix} y_1(i) \\ y_2^L(i) \\ y_2^R(i-1) \end{bmatrix} = \begin{bmatrix} 1 & \rho_{12} & \rho_{21} \\ \rho_{12} & e_2^L & 0 \\ \rho_{21} & 0 & e_2^R \end{bmatrix} \begin{bmatrix} b_1(i) \\ b_2(i) \\ b_2(i-1) \end{bmatrix} + \begin{bmatrix} n_1(i) \\ n_2^L(i) \\ n_2^R(i-1) \end{bmatrix} \qquad \text{(I.3)}$$

where

$$n_2^R(i) = \int_{(i+1)T+\tau_1}^{(i+1)T+\tau_2} n(t)s_2(t-iT-\tau_2)\,dt \qquad \text{(I.4a)}$$

$$n_2^L(i) = \int_{iT+\tau_2}^{(i+1)T+\tau_1} n(t)s_2(t-iT-\tau_2)\,dt \qquad \text{(I.4b)}$$

The channel in (I.3) has memory because of the dependence on previous inputs; however, since the random process $\{n(t)\}$ is white, the noise sequence in (I.3) in independent. Because of the sufficiency of the respective outputs, the channels in (2.4) and (I.3) have the same capacity region.

APPENDIX II

Proof of Lemma 1: To prove the identity

$$\det\left[I_{2n} + \sigma^{-2}E[X^nX^{nT}]R\right] = \det\left[I_{2n} + \frac{1}{\sigma^2}\begin{bmatrix}\Sigma_1 & 0\\ 0 & \Sigma_2\end{bmatrix}\begin{bmatrix} I_n & S^T \\ S & I_n\end{bmatrix}\right], \qquad \text{(II.1)}$$

we will first decompose the crosscorrelation matrices $E[X^nX^{nT}]$ and R using the Kronecker product:

$$E[X^nX^{nT}] = E\left[\left(X_1^n \otimes \begin{bmatrix}1\\0\end{bmatrix}\right)\left(X_1^n \otimes \begin{bmatrix}1\\0\end{bmatrix}\right)^T\right] + E\left[\left(X_2^n \otimes \begin{bmatrix}0\\1\end{bmatrix}\right)\left(X_2^n \otimes \begin{bmatrix}0\\1\end{bmatrix}\right)^T\right] = \Sigma_1 \otimes \begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix} + \Sigma_2 \otimes \begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix} \qquad \text{(II.2)}$$

and

$$R = I_n \otimes \begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix} + I_n \otimes \begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix} + S \otimes \begin{bmatrix}0 & 0\\ 1 & 0\end{bmatrix} + S^T \otimes \begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}. \qquad \text{(II.3)}$$

It is straightforward to check that if A, B, C, D are $n \times n$ matrices, then

$$A \otimes \begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix} + B \otimes \begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix} + C \otimes \begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix} + D \otimes \begin{bmatrix}0 & 0\\ 1 & 0\end{bmatrix} = P\begin{bmatrix}A & B\\ D & C\end{bmatrix}P^T \qquad \text{(II.4)}$$

where P is the permutation (orthogonal) matrix whose only nonzero entries are

$$P_{2j-1,j} = 1, \qquad j = 1, \cdots, n \qquad \text{(II.5a)}$$

$$P_{2j-2n,j} = 1, \qquad j = n+1, \cdots, 2n. \qquad \text{(II.5b)}$$

Therefore, (II.2) and (II.3) can be written, respectively, as

$$E[X^nX^{nT}] = P\begin{bmatrix}\Sigma_1 & 0\\ 0 & \Sigma_2\end{bmatrix}P^T \qquad \text{(II.6)}$$

$$R = P\begin{bmatrix}I_n & S^T\\ S & I_n\end{bmatrix}P^T \qquad \text{(II.7)}$$

and (II.1) follows from (II.6), (II.7), and the orthogonality of P upon using the identity $\det[I + AB] = \det[I + BA]$.

APPENDIX III

Proof of Lemma 2: Define the following diagonal matrices: $F = \operatorname{diag}\{\cos\theta_1, \cdots, \cos\theta_n\}$, $G = \operatorname{diag}\{e^{j\phi_1}\sin\theta_1, \cdots, e^{j\phi_n}\sin\theta_n\}$ where $\theta_i = (1/2)\arcsin(|\delta_i|) \in [0, \pi/4]$ and $\delta_i = |\delta_i|e^{j\phi_i}$, $i = 1, \cdots, n$. It is straightforward to check that

$$\begin{bmatrix}F & G^*\\ G & F\end{bmatrix}\begin{bmatrix}F & G^*\\ G & F\end{bmatrix}^* = \begin{bmatrix}F & G^*\\ G & F\end{bmatrix}^2 = \begin{bmatrix}I_n & \Delta^*\\ \Delta & I_n\end{bmatrix}. \qquad \text{(III.1)}$$

Therefore, we can write

$$\det\left|I_{2n}+\frac{1}{\sigma^2}\begin{bmatrix}A & 0\\ 0 & B\end{bmatrix}\begin{bmatrix}I_n & \Delta^*\\ \Delta & I_n\end{bmatrix}\right|$$
$$=\det\left|I_{2n}+\frac{1}{\sigma^2}P\begin{bmatrix}F & G^*\\ G & F\end{bmatrix}P^TP\begin{bmatrix}A & 0\\ 0 & B\end{bmatrix}P^TP\begin{bmatrix}F & G^*\\ G & F\end{bmatrix}P^T\right| \quad \text{(III.2)}$$

where P is the orthogonal matrix introduced in Appendix II. It follows from (II.4) that

$$P\begin{bmatrix}F & G^*\\ G & F\end{bmatrix}P^T$$

is a block diagonal matrix whose ith diagonal block is the 2×2 matrix

$$\begin{bmatrix}\cos\theta_i & e^{-j\Phi_i}\sin\theta_i\\ e^{j\Phi_i}\sin\theta_i & \cos\theta_i\end{bmatrix}$$

whereas

$$P\begin{bmatrix}A & 0\\ 0 & B\end{bmatrix}P^T$$

is a block matrix whose ith diagonal block is

$$\begin{bmatrix}a_{ii} & 0\\ 0 & b_{ii}\end{bmatrix}.$$

Because of the assumption on the magnitude of the elements of Δ, the hermitian matrix

$$\begin{bmatrix}I_n & \Delta^*\\ \Delta & I_n\end{bmatrix}$$

is nonnegative definite and, therefore, so is the matrix in the right side of (III.2). This allows us to apply the *generalized* Hadamard inequality [18][9] (the determinant of a positive-definite matrix is upper-bounded by the product of the determinants of its diagonal blocks) to the matrix in the right side of (III.2); the result is the inequality

$$\det\left|I_{2n}+\frac{1}{\sigma^2}\begin{bmatrix}A & 0\\ 0 & B\end{bmatrix}\begin{bmatrix}I_n & \Delta^*\\ \Delta & I_n\end{bmatrix}\right|$$
$$\le\prod_{i=1}^{n}\left\{\det\left[I_2+\frac{1}{\sigma^2}\begin{bmatrix}\cos\theta_i & e^{-j\Phi_i}\sin\theta_i\\ e^{j\Phi_i}\sin\theta_i & \cos\theta_i\end{bmatrix}\begin{bmatrix}a_{ii} & 0\\ 0 & b_{ii}\end{bmatrix}\begin{bmatrix}\cos\theta_i & e^{-j\Phi_i}\sin\theta_i\\ e^{j\Phi_i}\sin\theta_i & \cos\theta_i\end{bmatrix}\right]\right\}$$
$$=\prod_{i=1}^{n}\left\{1+a_{ii}+b_{ii}+a_{ii}b_{ii}(1-|\delta_i|)^2\right\}. \quad \text{(III.3)}$$

Finally, it is immediate to check that (3.31) is satisfied with equality if A and B are diagonal matrices.

APPENDIX IV

Proof of Lemma 3: The proof involves the solution of the maximization in (3.40), thus yielding an explicit way to compute the capacity region in (3.13). The maximum on the left side of 3.40) is achieved by the pair of nonnegative functions (Φ_1^*,Φ_2^*) that maximizes the strictly concave Lagrangian (if $\rho^2<1$, then $J(z_1,z_2,\rho^2)$ is strictly concave in (z_1,z_2))

$$L(\Phi_1,\Phi_2)=\int_{-\pi}^{\pi}f(\Phi_1(\omega),\Phi_2(\omega),\rho^2(\omega))\,d\omega$$
$$-\theta_1\left[\int_{-\pi}^{\pi}\Phi_1(\omega)\,d\omega-\frac{w_1}{\sigma^2}\right]-\theta_2\left[\int_{-\pi}^{\pi}\Phi_2(\omega)\,d\omega-\frac{w_2}{\sigma^2}\right] \quad \text{(IV.1)}$$

for some pair of positive Lagrange multipliers (θ_1,θ_2) selected so that

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\Phi_k^*(\omega)\,d\omega=\frac{w_k}{\sigma^2},\qquad k=1,2. \quad \text{(IV.2)}$$

The optimum pair (Φ_1^*,Φ_2^*) is the unique element in the cone of nonnegative function pairs that satisfies[10] (e.g. [33, p. 227])

$$\theta_1\ge f_1\left(\Phi_1^*(\omega),\Phi_2^*(\omega),\rho^2(\omega)\right)$$
$$=\frac{2\alpha-1}{1+\Phi_1^*(\omega)}$$
$$+\frac{(1-\alpha)\left[1+\Phi_2^*(\omega)\left(1-\rho^2(\omega)\right)\right]}{1+\Phi_1^*(\omega)+\Phi_2^*(\omega)+\Phi_1^*(\omega)\Phi_2^*(\omega)\left(1-\rho^2(\omega)\right)} \quad \text{(IV.3)}$$

$$\theta_2\ge f_2\left(\Phi_1^*(\omega),\Phi_2^*(\omega),\rho^2(\omega)\right)$$
$$=\frac{(1-\alpha)\left[1+\Phi_1^*(\omega)\left(1-\rho^2(\omega)\right)\right]}{1+\Phi_1^*(\omega)+\Phi_2^*(\omega)+\Phi_1^*(\omega)\Phi_2^*(\omega)\left(1-\rho^2(\omega)\right)} \quad \text{(IV.4)}$$

with equality if $\Phi_1^*(\omega)>0$ and $\Phi_2^*(\omega)>0$, respectively; here f_1 and f_2 are the partial derivatives of f with respect to its first two arguments. It follows from the second condition that

$$\Phi_2^*(\omega)=\max\left\{\frac{1-\alpha}{\theta_2}-\frac{1+\Phi_1^*(\omega)}{1+\Phi_1^*(\omega)\left(1-\rho^2(\omega)\right)},0\right\} \quad \text{(IV.5)}$$

which implies that

$$1-\alpha>\theta_2 \quad \text{(IV.6)}$$

for otherwise $\Phi_2^*(\omega)=0$ for all $\omega\in[-\pi,\pi]$, which does not satisfy (IV.2). Let us now see what conditions force each of the solutions to be zero at a particular frequency.

On the one hand, if $\Phi_1^*(\omega_0)=0$, then (IV.5) and (IV.6) imply that

$$\Phi_2^*(\omega_0)=\frac{1-\alpha}{\sigma^2}-1 \quad \text{(IV.7)}$$

which when substituted in (IV.3) results in

$$\rho^2(\omega_0)\ge\frac{\alpha-\theta_1}{(1-\alpha)-\theta_2}. \quad \text{(IV.8)}$$

One conclusion that can be drawn from condition (IV.8) is that

$$\alpha>\theta_1 \quad \text{(IV.9)}$$

because otherwise, $\Phi_1^*(\omega)=0$ for all $\omega\in[-\pi,\pi]$.

On the other hand, if $\Phi_2^*(\omega_0)=0$, then (IV.3) and (IV.9) imply that

$$\Phi_1^*(\omega_0)=\frac{\alpha}{\theta_1}-1 \quad \text{(IV.10)}$$

[9] Among the many existing proofs of the Hadamard inequality, Cover and El Gamal [32] have given a simple information-theoretic proof. One of the nice features of that proof is that it can be immediately extended to prove the generalized Hadamard inequality.

[10] If $|\rho_{12}|+|\rho_{21}|=1$, then $\rho^2(\omega_0)=1$ for either $\omega_0=0$ or $\omega_0=\pi$, and uniqueness of (Φ_1^*,Φ_2^*) is guaranteed, except in the set $\{\omega_0\}$ of measure zero, because $f(z_1,z_2,1)$ is not strictly concave.

which upon substitution in (IV.5) results in the condition

$$\rho^2(\omega_0) \geq \frac{\alpha}{1-\alpha}\frac{(1-\alpha)-\theta_2}{\alpha-\theta_1}. \tag{IV.11}$$

Note that since $1/2 \leq \alpha < 1$, (IV.8) and (IV.10) cannot be true simultaneously if $\rho^2(\omega_0) < 1$. However, they can indeed be false simultaneously, in which case (IV.3) and (IV.4) are satisfied with equality and $(\Phi_1^*(\omega), \Phi_2^*(\omega))$ are the positive solutions to the system of the following equations:

$$\Phi_2^*(\omega) = \frac{1-\alpha}{\theta_2} - \frac{1+\Phi_1^*(\omega)}{1+\Phi_1^*(\omega)\left(1-\rho^2(\omega)\right)} \tag{IV.12}$$

$$\frac{2\alpha-1}{1+\Phi_1^*(\omega)} + \frac{(1-\alpha)\left[1+\Phi_2^*(\omega)\left(1-\rho^2(\omega)\right)\right]}{1+\Phi_1^*(\omega)+\Phi_2^*(\omega)+\Phi_1^*(\omega)\Phi_2^*(\omega)\left(1r-\rho^2(\omega)\right)} = \theta_1. \tag{IV.13}$$

It follows that for each fixed pair of Lagrange multipliers, the maximizing spectra $(\Phi_1^*(\omega), \Phi_2^*(\omega))$ depend on ω only through $\rho^2(\omega)$, i.e., $\Phi_i^*(\omega) = \gamma_i(\rho^2(\omega), \theta_1, \theta_2)$, which is a continuous function of $\rho^2(\omega)$.

Appendix V

Proof of Lemma 4: Let E be the $2n \times 2n$ matrix whose only nonzero elements are $E_{1\,2n} = E_{2n\,1} = 1$, and let Z_i be the nonnegative-definite square root of Σ_i. Then we can write

$$\begin{aligned}\log\det&\left[I_{2n} + \frac{1}{\sigma^2}\begin{bmatrix}\Sigma_1 & 0\\ 0 & \Sigma_2\end{bmatrix}\begin{bmatrix}I_n & \tilde{S}^T\\ S & I_n\end{bmatrix}\right]\\ &- \log\det\left[I_{2n} + \frac{1}{\sigma^2}\begin{bmatrix}\Sigma_1 & 0\\ 0 & \Sigma_2\end{bmatrix}\begin{bmatrix}I_n & S^T\\ S & I_n\end{bmatrix}\right]\\ &= \log\det\left[I_{2n} + \rho_{21}ME\right]\end{aligned} \tag{V.1}$$

where

$$M = \sigma^{-2}\begin{bmatrix}Z_1 & 0\\ 0 & Z_2\end{bmatrix}\left[I_{2n} + \frac{1}{\sigma^2}\begin{bmatrix}\Sigma_1 & 0\\ 0 & \Sigma_2\end{bmatrix}\begin{bmatrix}I_n & S^T\\ S & I_n\end{bmatrix}\right]^{-1} \cdot \begin{bmatrix}Z_1 & 0\\ 0 & Z_2\end{bmatrix}. \tag{V.2}$$

The determinant in the right side of (V.1) is easily computed:

$$\begin{aligned}\det\left[I_{2n} + \rho_{21}ME\right] &= 1 + \rho_{21}^2 M_{12n}^2 + 2\rho_{21}M_{12n} - \rho_{12}^2 M_{11}M_{2n2n}\\ &\leq 1 + 2\rho_{21}M_{12n}\\ &\leq 1 + 2|\rho_{21}|\sqrt{M_{11}M_{2n2n}}\\ &\leq 1 + \frac{w_1+w_2}{\sigma^2}2n\end{aligned} \tag{V.3}$$

where the first two inequalities follow from the nonnegative definiteness of M, and the third inequality is a consequence of the fact that, according to (V.2), the largest eigenvalue of M is upper-bounded by that of

$$\sigma^{-2}\begin{bmatrix}\Sigma_1 & 0\\ 0 & \Sigma_2\end{bmatrix}$$

which is in turn upper-bounded by $(\operatorname{tr}\Sigma_1 + \operatorname{tr}\Sigma_2)/\sigma^2 \leq (w_1 + w_2)/\sigma^2$. Interchanging the roles of S and $\tilde{S}$, the same bound can be shown for the reverse difference in (V.1), i.e.,

$$\begin{aligned}\left|\frac{1}{n}\log\det\left[I_{2n} + \frac{1}{\sigma^2}\begin{bmatrix}\Sigma_1 & 0\\ 0 & \Sigma_2\end{bmatrix}\begin{bmatrix}I_n & \tilde{S}^T\\ S & I_n\end{bmatrix}\right]\right.&\\ \left. - \frac{1}{n}\log\det\left[I_{2n} + \frac{1}{\sigma^2}\begin{bmatrix}\Sigma_1 & 0\\ 0 & \Sigma_2\end{bmatrix}\begin{bmatrix}I_n & S^T\\ S & I_n\end{bmatrix}\right]\right|&\\ \leq \frac{1}{n}\log\left(1 + \frac{w_1+w_2}{\sigma^2}2n\right),&\end{aligned} \tag{V.4}$$

completing the proof of Lemma 4.

References

[1] R. Ahlswede, "Multi-way communication channels," in *Proc. 2nd Int. Symp. Information Theory*, Tsahkadsor, USSR, 1971, pp. 103–135.

[2] ——, "The capacity region of a channel with two senders and two receivers," *Ann. Prob.*, vol. 2, pp. 805–814, Oct. 1974.

[3] T. Cover, "Some advances in broadcast channels," *Advances in Communication Systems*, vol. 4. New York: Academic, 1975, pp. 229–260.

[4] A. D. Wyner, "Recent results in the Shannon theory," *IEEE Trans. Inform. Theory*, vol. IT-20, Jan. 1974.

[5] G. Sh. Poltyrev, "Coding in an asynchronous multiple-access channel," *Problems Inform. Transmission*, pp. 12–21, July–Sept. 1983.

[6] J. Y. N. Hui and P. A. Humblet, "The capacity region of the totally asynchronous multiple-access channels," *IEEE Trans. Inform. Theory*, vol. IT-31, pp. 207–216, Mar. 1985.

[7] S. Verdú, "Multiple-access channels with memory with and without frame synchronism," *IEEE Trans. Inform. Theory*, vol. IT-35, pp. 605–619, May 1989.

[8] M. A. Deaett and J. K. Wolf, "Some very simple codes for the nonsynchronized two-user multiple-access adder channel with binary inputs," *IEEE Trans. Inform. Theory*, vol. IT-24, pp. 63–64, Sept. 1978.

[9] J. Y. N. Hui, "Throughput analysis for code division multiple accessing of the spread spectrum channel," *IEEE J. Selected Areas Commun.*, vol. SAC-2, pp. 482–486, July 1984.

[10] G. Tartara, "Efficiency of channel use in asynchronous address systems with code address," *Problems Inform. Transmission*, vol. 6, pp. 144–147, Apr.–June 1970.

[11] R. C. Sommer, "Asynchronously multiplexed channel capacity," *Proc. IEEE*, vol. 54, pp. 79–80, Jan. 1966.

[12] H. L. Van Trees, *Detection, Estimation and Modulation Theory, I.* New York: Wiley, 1968.

[13] E. L. Lehmann, *Testing Statistical Hypotheses*. New York: Wiley, 1959.

[14] S. Verdú, "Capacity region of Gaussian CDMA channels: The symbol-synchronous case," in *Proc. 24th Allerton Conf.*, Oct. 1986, pp. 1025–1034.

[15] A. J. Viterbi and J. K. Omura, *Principles of Digital Communication and Coding*. New York: McGraw-Hill, 1979.

[16] U. Grenander and G. Szegö, *Toeplitz Forms and Their Applications*. New York: Chelsea, New York, 1958.

[17] B. Gyires, "Eigenwerte verallgemeinerter Toeplitzschen Matrizen," *Publ. Math. Debrecen*, vol. 4, pp. 171–179, 1956.

[18] R. Bellman, *Introduction to Matrix Analysis*. New York: McGraw-Hill, 1960.

[19] P. Lancaster and M. Tismenetsky, *The Theory of Matrices*. Orlando, FL: Academic, 1985.

[20] A. M. Geoffrion, "Strictly concave parametric programming, Part I: Basic theory," *Management Sci.*, vol. 13, pp. 244–253, Nov. 1966.

[21] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions*. Washington, DC: NBS, 1964.

[22] A. K. Jain, "A sinusoidal family of unitary transforms," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-1, pp. 356–365, Oct. 1979.

[23] R. M. Gray, "On the asymptotic eigenvalue distribution of Toeplitz matrices," *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 725–730, Nov. 1972.

[24] C. E. Shannon, "Communication in the presence of noise," *Proc. IRE*, vol. 37, pp. 10–21, Jan. 1949.

[25] R. G. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968.

[26] I. Csiszar and J. Korner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. New York: Academic, 1981.

[27] S. Verdú, "Minimum probability of error for asynchronous Gaussian multiple-access channels," *IEEE Trans. Inform. Theory*, vol. IT-32, pp. 85–96, Jan. 1986.

[28] R. G. Gallager, "A perspective on multiaccess channels," *IEEE Trans. Inform. Theory*, vol. IT-31, pp. 124–142, Mar. 1985.

[29] S. Gradshteyn and I. Ryzhik, *Table of Integrals, Series and Products*. New York: Academic, 1965.

[30] J. G. Proakis, *Digital Communications*. New York: McGraw Hill, 1983.

[31] F. D. Garber and M. B. Pursley, "Optimal phases of maximal sequences for asynchronous spread-spectrum multiplexing," *Electron. Lett.*, vol. 16, no. 19, pp. 756–757, Sept. 11, 1980.

[32] T. Cover and A. El Gamal, "An information-theoretic proof of Hadamard's inequality," *IEEE Trans. Inform. Theory*, vol. IT-29, pp. 930–931, Nov. 1983.

[33] D. G. Luenberger, *Optimization by Vector Space Methods*. New York: Wiley, 1969.

T-AC/34/9//29606

Control and Optimization Methods in Communication Network Problems

Anthony Ephremides
Sergio Verdú

Reprinted from
IEEE TRANSACTIONS ON AUTOMATIC CONTROL
Vol. 34, No. 9, September 1989

Control and Optimization Methods in Communication Network Problems

ANTHONY EPHREMIDES, FELLOW, IEEE, AND SERGIO VERDÚ, SENIOR MEMBER, IEEE

Abstract—In this paper we focus on two areas of communication network design in which methods of control and optimization theory have proven useful. These are the area of multiple access communication (for networks with shared links such as radio networks and local area networks) and the area of network routing (for networks with point-to-point interconnections). We review a few selected problems in each area to show the role of the control concepts involved and we then proceed to identify other areas of communication network design in which the same control theoretic and optimization methodology may be applicable and useful. We do not survey the work done in this area, nor do we review work in control areas whose methods are applicable in other communication network problems. Instead, we attempt to bring to the attention of the control systems community the numerous instances of problems arising in the pure communication network design process that can benefit from the attention and the capabilities of this community.

I. INTRODUCTION

COMMUNICATION networks are designed and built in order to share resources If interconnecting systems and bandwidths were available at no cost, then the solution to the problem of communication would be to assign dedicated communication links (channels) of sufficient capacity to every pair of conceivable users to meet their needs. This not being the case, it is necessary to multiplex the sources of communication traffic in order to optimize various cost criteria. Frequently, this optimization is dynamic and done on the basis of feedback that monitors the evolution of the degree of utilization of the network resources. Thus, we should expect a number of problems arising in communication network design to fit naturally in the framework of control systems design In this paper we wish to demonstrate that indeed this is the case and to show how various control and optimization methodologies have been used in the study of communication networks.

In the beginning there was a single communication network, the telephone network It represented a multibillion dollar investment and seemed to serve reasonably adequately the voice communication needs The explosive growth in data communication needs during the last 30 years built up the pressure for additional and alternative networking options. As a result, the notion of store-and-forward switching (known also as message switching) was introduced in the early 1960's. This notion represented a breakthrough since it constituted a radical reversal of thinking with respect to the circuit-switching process, namely, instead of securing an open, dedicated "pipe" for the transmission of messages by means of hardware switches, it allowed a step-by-step (node-by-node) forwarding of messages, thereby permitting each node to switch messages by deciding when and where to transmit the messages in its buffer. In the last 20 years we have seen an avalanche of technologies (fast switching, time division multiplexing, local area networks, fiber optical networks, integrated services digital networks, etc.) and a proliferation of operational public and private networks that put these technologies to test and challenged communication engineers. In addition, they should challenge control engineers as well.

Without attempting a survey of this vast application area we wish to promulgate the viewpoint that many (if not most) specific sub-problems in the network design process are natural control problems. In support of this thesis, we choose, first, to demonstrate how two major areas in communication networks (routing and multiple access) have benefitted from the use of techniques borrowed from what is traditionally perceived as control systems methodology and, second, to mention additional areas that are likely to benefit from the control systems community. As illustrated in this paper, the techniques that have proved useful in communication networks include. dynamic programming (e.g., [2], [6], [8]-[10], [22], [29], [38], [39], [47], [49], [54]); linear programming (e.g., [50], [51]); constrained and iterative optimization (e.g., [5], [14], [16], [42]); Markov decision theory tools (e.g., [2], [26], [29], [38]); control of Markov chains (e.g., [11], [17], [18], [20], [40], [45]); stability analysis of stochastic systems via Lyapunov methods (e.g., [31], [43]); sample path dominance (e.g., [2], [52]); and convergence of distributed and asynchronous algorithms (e.g., [6], [16], [42]).

The problem of routing is encountered in all and every network that does not permit the source to reach the destination in a single transmission hop, but instead it must traverse a path of intermediate links. By contrast, the problem of multiple access is encountered primarily in those networks that permit the nodes to reach their destination directly in one hop by having to share the same link with other transmitting nodes. In addition, the two problems are fundamentally different in nature and, jointly, cover considerable ground in the networking area. Finally, together they facilitate the identification of additional design issues and the extension of the applicability of suitable control methods. Thus, they represent "cornerstone" areas of network design.

Routing can be studied either macroscopically or microscopically. The macroscopic viewpoint considers basically a flow model and determines the splitting of the flow in order to reach the destination in minimum time with efficient use of the network resources. It is traditionally referred to as *static routing*. The microscopic viewpoint dissects the flow process down to the atomic level of the individual transmission unit, the message (a string of bits commonly referred to as packet), and determines the path each message must follow at each of its hops through the network. It is traditionally referred to as *dynamic routing*. Both viewpoints are explored in Section II.

Multiple access is a collective term that refers to numerous problems that deal with the dynamic allocation of a single resource among users who can coordinate their use of that resource only by making use of that resource. These problems arise primarily in the context of radio channels but also in the

Manuscript received September 13, 1988; revised December 27, 1988 and April 24, 1989. Paper recommended by Associate Editor, T. L. Johnson. This work was supported in part by the National Science Foundation under Grants NSF-85-00108 and NSF ECS 88 57689, by the Office of Naval Research under Contracts N00014-84-K-0207 and N00014-87-K-0054, and by the Army Research Office under Contract DAAL-03-87-K-0062.

A. Ephremides is with the Department of Electrical Engineering, University of Maryland, College Park, MD 20742.

S Verdú is with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544.

IEEE Log Number 8929606.

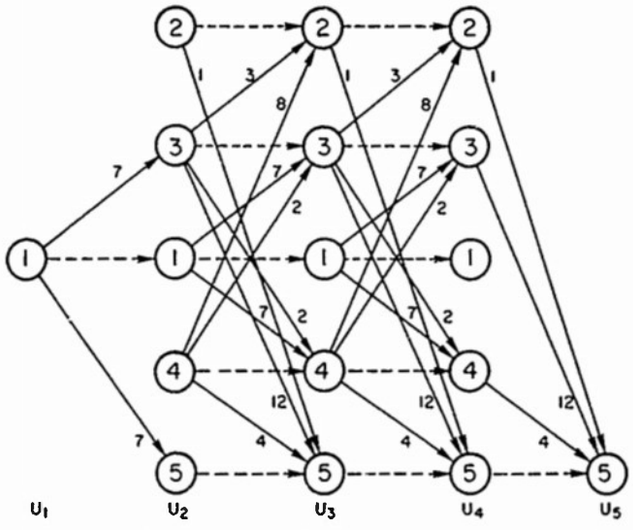

Fig 1 Layered network showing link lengths. Source is node 1 in U_1 and destination is node 5 in U_5.

context of shared cable resources in local area networks. In Section III, we explore the main multiple access problems where control methods have been successfully applied.

Both in the case of routing as well as in the case of multiple access we place the emphasis on the control techniques that have been used. We then show how these techniques, sometimes with slight modification, can be naturally transported to other problem areas such as voice-data integration, flow control, and the scheduling of messages and links. This is done in Section IV.

II. Network Routing

The problem of routing in communication networks is one that has received early attention and has experienced significant breakthroughs in the brief history of the field of communication networks. It is one of the first problems that gained prominence as a result of the emergence of store-and-forward switching. It is also one in which analytical tools and available theories applied nicely from the beginning.

A. Static Routing

Given a network (a set of nodes connected by directed links) a path connecting the source node to the destination node has to be selected from the set of all possible such paths.[1] In the simplest formulation, the problem is one of finding the shortest path, i.e., a length is assigned to each link and the optimization criterion is the total path length. This problem is one of the archetypical combinatorial optimization problems (the solution can be found by exhaustive enumeration of a finite set of possibilities—all possible paths from source to destination). Among the many existing shortest path algorithms (see, e.g., [41]), the Bellman-Ford algorithm (1956) is of particular interest to our exposition, both because it is based on dynamic programming and because, as we will see below, it easily lends itself to distributed asynchronous implementation. A natural choice to find the shortest path from source to destination in a layered network (i.e., one in which the nodes can be grouped in subsets $U_1, \cdots U_M$ such that the source and destination nodes belong to U_1 and U_M, respectively, and there are links only between nodes in adjacent layers U_{k-1} and U_k) such as the one in Fig. 1, is the dynamic programming algorithm, where the shortest paths and distances (costs-to-go) of the nodes in layer U_k to the destination are computed based on the shortest paths and distances of the nodes in layer U_{k+1}. If the network is not layered (such as that in Fig. 2), its shortest path can be obtained by finding the shortest path in a layered network derived from the original one as specified in the Bellman-Ford algorithm: the number of layers is equal to the number of nodes in the original network, say N, each layer contains a copy of each of the N nodes, and there is a link connecting two nodes in consecutive layers if such a link exists in the original network, in addition, copies of the same node in consecutive stages are connected by a zero-length link. (Fig. 1 was actually derived from Fig 2 using this rule) It is easy to see by induction that $D_k(i)$, the cost-to-go of node i in layer $N - k$, is the minimum length of any path from i to the destination that uses at most k links (in the original network). Since no shortest path uses more than $N - 1$ links (link lengths are assumed nonnegative and, therefore, no path containing loops need be considered), the cost-to-go of node i at layer 1, $D_{N-1}(i)$ will indeed be the length of the shortest path from node i to the destination. Thus, the Bellman-Ford algorithm can be formulated as the iteration

$$D_k(i) = \min_{j \in N(i)} [D_{k-1}(j) + d_{ij}] \qquad \text{for } k = 1, \cdots N-1 \quad (2.1)$$

where d_{ij} is the length of the link from i to j, $N(i)$ is the set of nodes for which such a link exists and it is assumed that $D_0(i) = \infty$ if i is not the destination node, which corresponds to the removal of all the nodes but the destination in the final layer (Fig. 1).

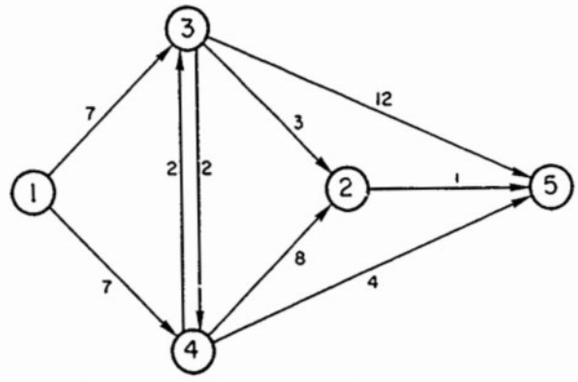

Fig. 2. Arbitrary network showing link lengths. Source is node 1 and destination is node 5.

Contrary to what may appear at first glance there is a lot more to network routing than finding shortest paths. After all, the shortest path may not be the best path. The reason is that the real goal is to minimize the *delay* experienced in going from source to destination, and the delay encountered in each link is usually a function of the amount of traffic carried by the link (as the link becomes congested, it takes longer to go through it), which is referred to as the link flow and is quantified in packets (or messages) per second. Then, assuming a given desired flow level from source to destination, the problem is how to distribute it among all the possible paths so as to minimize the total delay. In contrast to the previous more elementary formulation of the routing problem which led to the shortest path combinatorial optimization problem and which corresponds to the special case in which the link delays are independent of the flows, we now face a continuous optimization problem which can be written as

$$\text{minimize } F(x) = \sum_{(i,j)} D_{ij}\left(\sum_{n \in P(i,j)} x(n)\right)$$

$$\text{subject to } x \in X = \left\{(x(1), \cdots x(J)) \in R^J, \sum_{n=1}^{J} x(n) = \lambda,\ x(n) \geq 0\right\} \quad (2.2)$$

[1] All the algorithms and results discussed in this section can be extended to the case where there are several source-destination pairs in the network.

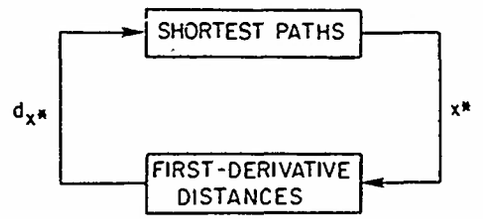

Fig 3 Characterization of the solution to the minimum-delay routing problem.

where the set of all paths from source to destination is labeled $\{1, \cdots, J\}$; $x = (x(1), \cdots, x(J))$ is the vector of unknown nonnegative path flows which sum up to λ, the desired flow from source to destination; $P(i,j) \subset \{1, \cdots, J\}$ is the subset of paths that traverse link (i, j); and $D_{ij}(x)$ is the portion of the overall delay contributed by the link from node i to node j when the flow it carries is equal to x. In order to characterize a global solution to the optimization over a convex set in (2.2), it is natural to restrict attention to convex penalty functions. In practice, it is common that the incremental delay in a link grows with the amount of traffic it carries and, therefore, it can be assumed that the functions D_{ij} are convex without affecting significantly the practical applicability of the results.

Now, the characterization of the solution to (2.2), x^*, is straightforward. Since the feasible set X and the penalty function F are convex, it is necessary and sufficient that the directional derivative of the penalty function be nonnegative when evaluated at x^* in the direction of any of the elements of X (e.g., [37])

$$0 \le \lim_{\alpha \downarrow 0} \frac{1}{\alpha} [F((1-\alpha)x^* + \alpha x) - F(x^*)] \qquad \forall x \in X \quad (2.3)$$

which translates into

$$0 \le \sum_{(i,j)} D'_{ij} \left(\sum_{m \in P(i,j)} x^*(m) \right) \sum_{n \in P(i,j)} [x(n) - x^*(n)]$$

$$= \sum_{n=1}^{J} [x(n) - x^*(n)]\, d_{x^*}(n) \qquad \text{for all } x \in X \quad (2.4)$$

where $d_x(n) = \Sigma_{(i,j) \in L(n)} D'_{ij}(\Sigma_{m \in P(i,j)} x^*(m))$ is the length of path n when the length of each link is equal to the derivative of its delay evaluated at the set of flows x, and $L(n)$ is the set of links used by path n. The solution to (2.4), x^*, is the vector in X that minimizes its inner product with the vector of distances d_{x^*}. Thus, x^* puts all its weight on the smallest component(s) of d_{x^*}. The conclusion is that the optimum flow uses only shortest paths computed according to the *derivative* of the link delays.

This solution to the minimum-delay routing problem allows us to check whether a given set of flows is optimum. Unfortunately, it does not tell us how to find the optimum flows. Indeed, we face the chicken-and-egg situation depicted in Fig. 3. The optimum flows are obtained by solving a shortest path problem; but in order to compute the link lengths it is necessary to know the optimum flows. Nevertheless, the foregoing characterization of the optimal solution does suggest a possible iterative procedure to find the optimum set of flows. Starting with a given set of flows x one can compute the minimum derivative shortest paths for that flow, and hence, a new flow, $x^*(x)$ that is positive only along those shortest paths. The process can then be repeated, until there is no appreciable cost decrease. The region of convergence of such a procedure can be improved by letting the new flow be a convex combination of x and $x^*(x)$, i.e.,

$$x_{k+1} = (1-\alpha_k)x_k + \alpha_k x^*(x_k).$$

This is the so-called *flow deviation method* of Fratta, Gerla, and Kleinrock [14], where $0 \le \alpha_k \le 1$ is chosen to minimize

$$F((1-\alpha_k)x_k + \alpha_k x^*(x_k))$$

which is a special case of the feasible-direction nonlinear programming algorithm due to Frank and Wolfe [13]. The convergence of the flow deviation method to the optimum routing is rather slow because unfavorable paths tend to carry considerable flow during many iterations unless the initial routing guess is particularly fortuitous. Such a behavior can be improved by reducing the flow along each nonminimum derivative path in accordance to the delay experienced in that path. This is the idea of iterative routing algorithms based on *gradient projection* nonlinear optimization methods (e.g., [4]) in which the flow decrease along a nonminimum derivative path is proportional to the difference between its length and that of the shortest path (according to the first derivative of the delay function). If such a decrease would result in a negative flow, then the flow along that path is set to zero (hence, the *projection* to the set of feasible flows).

We have seen that the problem of static network routing can be formulated as a conceptually straightforward optimization problem that admits well-known solutions in nonlinear programming. What sets optimum routing in communication networks apart from other multicommodity flow problems arising in operations research is the fact that the optimization is carried out in real time, and often, in distributed fashion, where each node makes its own routing decisions based on local information. The review of centralized routing has revealed that the shortest path problem plays a central role in solving for the optimum routing regardless of whether the link congestion measures depend on the link flow or not. Hence, we will start the exposition of distributed routing algorithms by discussing the distributed version of the Bellman-Ford shortest path algorithm.

The Bellman-Ford updating equation in (2.1) suggests that the algorithm is suited for decentralized operation because each node can update its own estimate of distance to the destination (cost-to-go) provided it receives from its neighbors their own estimates [appearing on the right-hand side of (2.1)]. The feature that makes the study of the distributed Bellman-Ford algorithm interesting is that it can run completely *asynchronously*, in the sense that the updating and communication times need not be coordinated and convergence can be guaranteed by simply assuming that updating and communication between nodes never cease, without any requirements whatsoever on the rate of communication. The proof of convergence is a nice illustration of the analysis of decentralized algorithms where the processors are allowed to perform their computations and to communicate the corresponding results completely independently of one another [5], [6]. The idea is to show that the estimates computed in the distributed asynchronous algorithm are always sandwiched by the estimates computed by the centralized version of the algorithm when started at two different initial conditions, and that both centralized estimates converge to the true distances to the destination node.

Those centralized estimates are denoted by $\bar{D}_k = (\bar{D}_k(1), \cdots, \bar{D}_k(N))$ and $\underline{D}_k = (\underline{D}_k(1), \cdots, \underline{D}_k(N))$, and are the result of the centralized Bellman-Ford iteration (2.1) when it is started with initial conditions $\bar{D}_0 = (\infty, \cdots, \infty, 0)$ and $\underline{D}_0 = (0, \cdots, 0)$, respectively. (The destination node is assumed to be the Nth node.) Define the operator [see (2.1)]

$$B_i[D_k] = \min_{j \in N(i)} [D_k(j) + d_{ij}]$$

$$= D_{k+1}(i) \quad (2.6)$$

if $1 \le i < N$, and $B_N[D_k] = D_k(N)$. This operator is monotone in the sense that if $D \le D^*$ (i.e., if $D(i) \le D^*(i)$, $i = 1, \cdots N$), then

$$B_i[D] \le B_i[D^*]. \quad (2.7)$$

The monotonicity of B_i implies that

$$\underline{D}_k \le \underline{D}_{k+1} \le \bar{D}_{k+1} \le \bar{D}_k \quad (2.8)$$

and, moreover, it is easy to show that for sufficiently large k

$$\underline{D}_k = \bar{D}_{N-1} = \bar{D}_k \tag{2.9}$$

which is the vector of distances from each node to destination as we saw in the discussion of the centralized algorithm.

In the asynchronous distributed version of the algorithm, it is assumed that each node i keeps at time $t \geq 0$ an estimate of its distance to destination $A_t(i)$, and an estimate of the distance from each of its neighbors $j \in N(i)$ to destination $A_t^i(j)$, which is simply the latest estimate received from node j. In view of (2.8) and (2.9), convergence of the algorithm will follow if we show that for every index k, there exists a time $t_k > 0$ such that for all $t \geq t_k$

$$\underline{D}_k \leq A_t \leq \bar{D}_k \tag{2.10}$$

and for $i = 1, \cdots, N-1$

$$\underline{D}_k(j) \leq A_t^i(j) \leq \bar{D}_k(j) \qquad j \in N(i). \tag{2.11}$$

This is shown by induction. If $k = 0$, then (2.10) and (2.11) hold as long as the initial estimates of the decentralized algorithm are nonnegative. Assuming that the induction hypothesis is true for k, the monotonicity of B_i implies that if $t \geq t_k$, then

$$\underline{D}_{k+1}(i) = B_i[\underline{D}_k] \leq B_i[A_t^i] \leq B_i[\bar{D}_k] = \bar{D}_{k+1}(i). \tag{2.12}$$

But $A_t(i)$ is a piecewise constant function of time which only jumps at the updating times of node i, at which times it takes the value

$$A_t(i) = B_i[A_t^i].$$

Therefore, we can write

$$\underline{D}_{k+1}(i) \leq A_t(i) \leq \bar{D}_{k+1}(i) \qquad \text{for } t \geq t_k(i) \tag{2.13}$$

where $t_k(i)$ is the smallest updating time of node i which is greater than t_k. Moreover, if we wait long enough after $\max_i t_k(i)$, not only all the nodes will have carried out their first updates after t_k but the result of those computations will have been communicated to their neighbors because of the assumption that updating and communication occur infinitely often. Hence, there exists $t_{k+1} \geq \max_i t_k(i)$ such that for all $t \geq t_{k+1}$ and for all i and j

$$A_t^i(j) = A_s(j)$$

for some $s \geq t_k(j)$ (which depends on t, i, and j). Thus, it follows from (2.13) that

$$\underline{D}_{k+1}(j) \leq A_t^i(j) \leq \bar{D}_{k+1}(j) \qquad j \in N(i) \qquad i = 1, \cdots, N-1$$

completing the induction proof and, therefore, the proof of convergence of the distributed asynchronous Bellman-Ford algorithm.

When the link delays depend on the traffic flows, it is also possible to obtain the optimal routing that solves (2.2) in a distributed asynchronous fashion. Gradient projection algorithms are better suited for this task than the flow deviation method because in the latter method a higher degree of synchronization is required in order for the nodes to use the same step size at each iteration. In the distributed asynchronous implementation of gradient projection optimum routing algorithms, each node broadcasts from time to time the values of its outgoing flows to its upstream neighbors, who in turn pass that information on to their upstream neighbors. In this way, the source keeps estimates at all times of the link flows and can carry out the gradient projection iteration autonomously based on those estimates. The first algorithm based on this idea was due to Gallager [16], who posed an alternative formulation to (2.2), where the unknowns are the fractions of flow routed to each outgoing link at each node, rather than the path flows. Tsitsiklis and Bertsekas [42] showed the convergence of the distributed asynchronous implementation of gradient projection optimal routing algorithms provided the time between consecutive broadcasts is small enough relative to the speed at which the flows generated by the algorithm change. The approach for showing the stability of this algorithm is very different from the proof of convergence of the distributed Bellman–Ford algorithm where the monotonicity of the dynamic programming mapping implied that the estimates are closer and closer to the solution regardless of the actual sequence of communication and computation times. The idea here is that if the step size of the algorithm is small enough, then the flows change so slowly with respect to the periods between communication times that their evolution is very close to that of the centralized algorithm which uses the unique, true value of each link flow

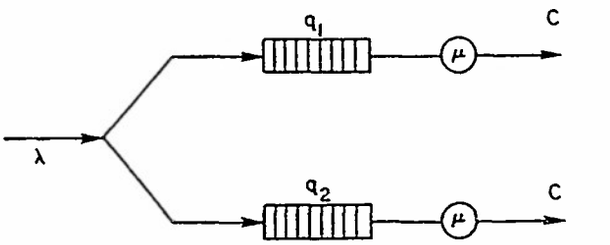

Fig. 4. Queueing model of a node with one incoming link and two outgoing links.

B. Dynamic Routing

As mentioned earlier, there are two fundamentally different philosophies to network routing: either viewing it as a "flow" problem in which the traffic of messages is modeled as a "macro"-commodity entering the network as a single entity (static or quasi-static routing), or as an individualized-message path-finding problem in which the traffic is broken down to its constituent elementary units (dynamic routing)—a dichotomy akin to that of statistical/quantum mechanics in physics. Whereas the first approach leads to optimization problems where time plays no role, the essential ingredient of the second approach is the randomness of the time-evolution of the buffers in the network, thus placing dynamic routing within the sphere of stochastic control.

• The most elementary instance of dynamic routing is the simple queueing system shown in Fig. 4 which models a node with one incoming link and two outgoing links. It simplifies considerably the dynamics of the message arrival process and of the service time characteristics and ignores processing delay. Thus, the arrival instants of messages over the incoming link are assumed to constitute a Poisson process of constant rate λ. Upon arrival each message is put in the buffer of one of the two outgoing links This action represents the "control." The buffers are assumed to have unlimited (infinite) capacity and the message lengths are assumed to be random with exponential distribution (an obvious additional simplification) with parameter μ. The two outgoing links have equal capacity of C bits/s. Thus, each link is modeled as a queueing system with exponential service time distribution with parameter μC. It is desired to characterize the optimal control policy that minimizes the average total delay per message based on the observations of the "state" of the system, namely the number of messages q_1 and q_2 in the two buffers. The model, of course, assumes that the head-of-the-line message is dropped from the buffer as soon as the transmission of its last bit is completed

This model, despite its simplicity, proved to be rather difficult to analyze. For details, see [10], it is not important to repeat them here. It should suffice to state that the main result, which simply requires that upon arrival a message should join the shortest queue (with arbitrary decision in case the two queues have equal numbers of messages), was hardly surprising Yet an intricate

argument on the dynamic programming equation (DPE) was needed and there were some counter-intuitive side-results including the relaxation of the Poisson assumption on the arrivals, and the fact that in the incomplete state information case, the *certainty-equivalent control* (i.e., send the message to the *expected* shortest queue) need not be optimum unless both queues have the same number of customers initially.

The optimality of the send-to-shortest-queue (SS) policy in the complete state information case can be proved in a rather strong sense. At all times, the sum $(q_1 + q_2)$ and maximum $(\max\{q_1, q_2\})$ of the number of messages in both buffers are stochastically minimized by the SS policy in the sense of the partial order between random variables according to which the random variable X is *stochastically smaller* than Y if $P[X \le a] \ge P[Y \le a]$ for all a. The proof of optimality can be obtained by the method of *forward induction* [53], whereby the desired stochastic ordering between the queue sizes under the optimum and an arbitrary policy is shown to be preserved at each transition.

The problem formulation of [10] is one of many related ones (see [8], [9], [22], [24], [33], [38], [54], [55]) which are slightly more complicated but share some fundamental characteristics which, in fact, extend beyond the confines of the routing problem into the areas of priority assignment, resource allocation, and flow control. They are all Markovian decision process (MDP) problems. In the sequel we will describe a fairly general MDP that includes the dynamic routing problem as a special case. In fact, it includes almost all of the queueing control problems that have been studied in connection with communication network issues. We will then outline the solution methodologies that have been used. These include basically: 1) the derivation of optimality conditions from the DPE associated with the corresponding MDP; 2) the use of sample path stochastic dominance arguments, and finally; 3) the reformulation of the MDP as a linear program. We should emphasize, lest the reader be unduly encouraged, that the problems in this area are sufficiently complex, so that only modest results can be generally obtained despite involved arguments and nontrivial machinery. Typically, these results characterize some structural properties of the optimal policy. However, knowledge of such structure is often sufficient to permit close approximation of the actual optimal policy by well-founded heuristics.

Let us recall briefly what an MDP is (for details, see [30]). We need a state description of the process to be controlled. Let S be its state space. When in state $s \in S$, a set A_s of admissible control actions is specified. When action $a \in A_s$ is applied, there is a transition from state s to s' that is governed by the probability distribution $p(s'|s, a)$, and which occurs after a random time τ which is exponentially distributed with distribution denoted by $(\tau|s, a, s')$. Clearly, p and t together describe the stochastic dynamics of the process to be controlled. Finally, each transition is accompanied by a cost penalty that we denote by $c(\tau, s, a, s')$.

The dynamic routing problem we considered before fits in this formulation easily. In that case, the state space is $S = \{0, 1, 2, 3, \cdots\}^2$. An element $s = (q_1, q_2) \in S$ is simply the pair of values of the respective queue sizes. The set of actions A_s is the same for any state and consists of a_1 and a_2 where a_i is the action that assigns an arriving message to the buffer of link i. The distribution p is of trivial form, in that the transitions are deterministic. Assignment of an arrival to queue i augments q_i by one. Note, now, that in addition to the arrival instants, the departure (or service completion) instants are important because they induce state transitions as well. A departure from queue i reduces q_i by one. When a departure occurs there is no meaningful control action that can be applied in this particular problem. The exponential distribution t corresponds to times between arrivals and/or departures.[2] Finally, the cost rate c must reflect the delay. By Little's result in queueing theory, we know that the average delay is proportional to the average number of customers in the queue. Thus, $c(\tau, a, s, s')$ can be taken to be simply equal to $(q_1 + q_2)$. This MDP formulation can be extended to encompass more complicated queueing control problems.

Let us return now to the general MDP. We need to specify the notion of a control policy and the optimization criterion. Let us denote by $\xi_1, \xi_2, \cdots$, the state transitions that occur at instants $t_1, t_2, \cdots$. A policy π is a sequence of decision rules $\pi_1, \pi_2, \cdots$, where π_n determines the choice of action at the transition time t_n. It can be viewed as a conditional distribution on the set of actions parametrized by the past history of the process.

The optimization criterion that corresponds to the practical case of expected total delay is the long-run average expected cost; namely, if we denote by $V(\pi, i, t)$ the expected cost incurred under policy π, with initial state i, until time t we consider as the optimization criterion the value function

$$V(\pi, i) \triangleq \liminf_{t\to\infty} \frac{V(\pi, i, t)}{t}.$$

For technical reasons, however, that are well known to optimization specialists, it is easier to establish optimality conditions if we consider, instead, the so-called α-discounted cost, i.e.,

$$V^\alpha(\pi, i) = \int_{t=0}^{\infty} e^{-\alpha t}\, dV(\pi, i, t).$$

The latter converges to the former as $\alpha \to 1$ under a variety of stationarity conditions. For technical reasons that will become apparent in the sequel, we will also consider the finite-horizon costs. These are defined in a similar fashion except that we let time extend only to t_n, the instant of the nth transition. If we denote by $V^\alpha(i)$ and $V(i)$ (and also $V_n^\alpha(i)$, $V_n(i)$ for the finite horizon cases) the values of these cost functions when π is chosen optimally, we are led to the following DPE:

$$V^\alpha(i) = \inf_{a\in A_i} \sum_{i'} [c(i, a, i') + \beta(i, a, i') V^\alpha(i')] p(i'|a, i)$$

or

$$V_{n+1}^\alpha(i) = \inf_{a\in A_i} \sum_{i'} [c(i, a, i') + \beta(i, a, i') V_n^\alpha(i')] p(i'|a, i)$$

where

$$\beta(s, a, s') \triangleq \int_0^\infty e^{-\alpha\tau}\, dt(\tau|s, a, s')$$

and

$$c(s, a, s') \triangleq \int_0^\infty c(\tau, s, a, s')\, dt\,(\tau|s, a, s')$$

are the discount factor and cost values per transition, respectively.

The DPE is of fundamental importance in the study of MDP's because the value function V^α has the usually convenient properties of convexity, supermodularity, and other forms of monotonicity that lead readily to sufficient conditions for optimality. The difficulty with the analysis of the DPE is that the optimality conditions are heavily problem-dependent and often lead to explosively large numbers of cases to be verified separately. This is especially true for MDP's that arise from queueing models. For this reason, and because of additional difficulties that arise when the state is on the boundaries (see [22]), it became evident that alternative methods of solution were needed.

[2] A slight modification of the model of transitions, called uniformization, is useful in that it introduces dummy transitions from a state into itself, thus some situations which introduce nonessential complications can be handled without departure from this discrete transition time formulation.

One alternative method that has received attention recently and which produced successful results in problems of queueing control (akin to the routing problem) is a probabilistic method called sample-path or stochastic dominance. This method bypasses completely dealing with the value function. Instead, it focuses directly on seeking the optimal policy. Let G be the class of admissible policies. If we suspect that the optimal policy π has a property p, then we can proceed as follows in order to prove that it actually does have that property. Let S be a subset of G, to which we know the optimal policy belongs. We consider a subset of policies $S_p \subset S$, all elements of which have the property p. For every $\pi \notin S_p$, we attempt to construct a policy $\hat{\pi}$ which outperforms π. If we succeed, we must conclude that the optimal policy belongs to S_p. In constructing $\hat{\pi}$ we often need to engage in a careful reorganization of the underlying probability space in order to align the sample paths properly, so that the comparison of the two policies can be made for every sample path. This procedure is full of risks and extreme care is required to avoid faulty arguments. Note, also, that to apply this method usefully, we must have "guessed" the properties of the optimal policy correctly. Thus, at best, it is a method to verify the validity of our conclusions, rather than a method that leads us to the right conclusions.

Successful use of the stochastic dominance approach was made in [52] and [50] where a problem that is dual to the problem of dynamic routing was studied. Specifically, in a two-server queueing system in which the two servers have unequal service rates, we wish to determine whether and when the slower server needs to be activated if we are interested in minimizing the usual total expected delay function. That the optimal policy has a threshold form (namely that the slower server must be activated when the queue size exceeds a crucial value) was proven in [29] via the DPE method. However, the alternative proof via the arguments of stochastic dominance was much simpler and led to a generalization of the result to cases of nonexponential arrivals and/or service, that could not have been easily accomplished by means of the DPE method.

Another successful use of the stochastic dominance method has been noted in [2]. In this case the problem of optimally choosing which customer to serve next in a single queueing system was considered under the constraint that each customer must begin (or terminate) service by an individually assigned random deadline or else it is dropped from the system. The cost criterion is then to minimize the expected number of lost customers. It was proven that scheduling the customer with shortest time to extinction minimizes this cost.

Although these problems differ from routing, the model structures are quite similar, and it has been observed that, usually, queueing control problems with such structural similarities can be studied equally successfully.

The third method, which was first used in [38] in the study of a specific queueing control problem, and which has been broadly extended recently in [51], is the linear programming approach. Almost any queueing control problem that can be formulated as a MDP (therefore the problem of dynamic routing, as well) can be converted to an equivalent linear program (LP). The advantages of this conversion are that it is problem-independent and it leads occasionally to successful study of semi-Markov decision problems as well. Furthermore, it facilitates considerably the characterization of optimal solution properties. Here is how this equivalence can be demonstrated.

Let us concentrate on an MDP under a finite-horizon, discounted cost formulation.[3] We shall consider a queueing model with state dynamics given by

$$x_{k+1} = x_k + \xi_{k+1} z_{k+1}.$$

Here, x_k denotes the state at t_k (the instant of the kth transition), ξ_k represents that transition, and z_k represents the control action at that transition. The transition ξ_k can represent an arrival or a departure as an increment of the state. The control z_k is conveniently defined to enable ($z_k = 1$) or disable ($z_k = 0$) a transition. For example, in the routing model discussed at the beginning of the section, the state is equal to a two-dimensional vector of queue sizes, and the transition corresponding to sending an arriving message to the first queue would be represented by $\xi_k = [1\ 0]^T$. Indeed, a variety of queueing control problems (in fact, the vast majority of those that have been considered in connection with communication network problems) can be so represented.

Note that the crucial aspect of this state equation is the *linear* dependence on the controls. Note also that usually the cost function is linear in the state (since the usual cost criterion is the expected delay which is coupled to the queue sizes, and hence the state, by Little's result). Consequently, the cost is linear in the controls. The minimization of the cost over the set of control trajectories is constrained since the state equation must be satisfied and the state must always belong to an admissible set (typically, a set of vectors with integer-valued coordinates belonging to given ranges). Thus, the constraints are also linear in the controls, and the problem is easily formulated as an LP. There are, however, two points that require attention. First, the controls are integer-valued, i.e., $z_k \in \{0, 1\}$. Second, in the MDP the vectors ξ_k are random and depend on past history.

The first problem is taken care of in one of two ways: by construction or by use of a property of the constraint matrix of the linear program, called unimodularity. The construction method involves using a noninteger optimum control whose quantized version satisfies the MDP optimality conditions (see [38], [51] for details). The use of unimodularity involves a well-known result in the theory of integer linear programming (e.g., [34]): if the constraint matrix of an LP is integer-valued and totally unimodular (i.e., each of its sub-determinants is $+1$, -1, or 0), then all the vertices of the feasible polytope are integer-valued. Therefore, no further restrictions are needed to guarantee that the solution of a conventional LP will result in the integer-valued optimal control. Fortunately, in many queueing problems of interest (including the dynamic routing problem), the constraint matrix is indeed totally unimodular.

The second problem is easily taken care of by thinking of the z_k's as functions from the sample space Ω to the action space. Thus, the cost criterion can be written as a functional on the underlying probability space.

Let $z_k(\omega_k)$ represent the control action at the kth transition, where ω_k denotes the random "history" until the kth transition. We have

$$x_{k+1}(\omega_{k+1}) = x_k(\omega_k) + z_{k+1}(\omega_{k+1})\xi_{k+1}(\omega_{k+1}).$$

Let S and Z be the set of admissible states and controls, respectively. The β-discounted, n-step, expected cost under policy z and initial condition x is given by

$$J_n^{\beta}(x, z) = E_z \sum_{k=0}^{n-1} \beta^k L(z_k)$$

where

$$L(z_k) = c^T x_k + d^T z_k$$

(c and d denote constant column vectors). This is a cost function that is adequately general. For example, in a pure resource allocation problem without blocking or rejection of messages we have $d = 0$, while in pure blocking problems we take $c = 0$. The state equation, after repeated iterations, yields

$$x_k(\omega_k) = x + \sum_{j=1}^{k} z_j(\omega_j)\xi_j(\omega_j), \qquad k > 0.$$

[3] The reason that we cannot work directly with infinite horizons is the possibility of so-called duality gaps in linear programming theory with infinite-dimensional variables.

Therefore,

$$J_n^\beta(x, z)=E_x\sum_{k=0}^{n-1}\beta^k\left\{c^Tx+c^T\sum_{j=1}^{k}z_j\xi_j+d^Tz_k\right\}$$

$$=\frac{1-\beta^n}{1-\beta}c^Tx+E_x\sum_{k=1}^{n}\beta^k\left\{\sum_{j=1}^{k}c^Tz_j\xi_j+d^Tz_k\right\}.$$

But

$$E_x(z_k)=\sum_{\omega_k}z_k(\omega_k)Pr(\omega_k).$$

Hence

$$J_n^\beta(x, z)=\frac{1-\beta^n}{1-\beta}c^Tx+\sum_{k=1}^{n}\sum_{\omega_k}\gamma_k(\omega_k)z_k(\omega_k)$$

where $\gamma_k(\omega_k)$ is a known function that depends on $Pr(\omega_k)$, c, ξ_k, and β^k. Consequently, the MDP is equivalent to

$$\min_{z_k}\sum_{k=1}^{n}\sum_{\omega_k}\gamma_k(\omega_k)z_k(\omega_k)$$

subject to

$$\left(x+\sum_{j=1}^{k}z_j(\omega_j)\xi_j(\omega_j)\right)\in S$$

which is a conventional LP where the initial condition plays the role of a parameter, the sensitivity with respect to which can be studied by the well-developed theory of sensitivity analysis of linear programming [15].

In conclusion, we see that the MDP is converted to an equivalent LP under very mild conditions that are usually satisfied by dynamic routing and other queueing control problems. Thus, a third alternative methodology becomes generally available for the study of these problems. Whether to choose from the arsenal the DPE approach, or the LP method, or stochastic dominance tools, depends on the problem and on the, as yet undeveloped, intuition that the investigator should possess.

III. Multiple-Access Communications

The communication networks considered in the discussion of routing problems in Section II consist typically of a set of nodes connected by point to point communication links. Each of these links viewed in isolation can be modeled as a classical communication channel with one sender and one receiver. In this section, we consider multipoint to point communication links where several transmitters share a common channel. Multiple access channels are the basic building blocks of radio networks, satellite communication, and local area networks, and during the last 15 years have attracted the attention of many communication, information, and control theorists.

There is a wide variety of strategies to divide the "resource" of a communication channel among several geographically dispersed transmitters. The simplest methods are those that assign a permanent independent sub-channel to each transmitter (e.g., in frequency division multiple access and time division multiple access); these strategies are easy to analyze and are widely used in practice in situations where the users need to transmit at fairly steady rates. If the transmitters are *bursty* (i.e., the radio of peak-to-average rate at which the need to transmit is high) those static methods are inefficient since most of the time the channel is underutilized while demand (and induced delay) accumulates at busy terminal locations. Dynamic channel sharing strategies overcome this problem by allocating channel resources on an on-demand basis. Consistent with the overall spirit of this paper, our goal here is not to review this vast topic, but rather to demonstrate how control theory can play a useful role in its study. Here we wish to single out two multiple access strategies: random access and simultaneous transmission, which are broadly representative of dynamic channel sharing systems and in which control theoretic concepts have played a pivotal role.

In random access communication, the conceptual allocation model is addressed without an effort to exploit the signaling degrees of freedom and the micro-structure of the transmitted messages. For this purpose, a crude channel model is considered, that achieves this separation of the "macro" from the "micro" problem. In simultaneous transmission systems, however, a more refined viewpoint is adopted, by taking the realities of the medium into account, modeling them, and exploiting them.

A. Random-Access

The object of interest here is the so-called *collision channel* model, in which messages (called packets) require one time unit (called slot) for transmission and are sent by a population of users who are synchronized so that their slots coincide at the receiver, but are otherwise uncoordinated and unaware of which and how many users have packets to transmit. If two or more packets are simultaneously transmitted, it is assumed that the receiver is unable to recover any of the messages, and they have to be retransmitted in a future slot. In the ALOHA algorithm, which was developed in the early 1970's [1] at the University of Hawaii and marked the beginning of the area of random-access communication, each packet that has been unsuccessfully transmitted before is transmitted with probability p in the next slot. New packets which have not attempted transmission before are transmitted with probability either 1 or p depending on which version of the ALOHA algorithm is used. In our discussion, we will assume the latter choice.

Under these conditions, and assuming that the number of newly generated packets in each slot is a random variable (with mean λ) independent from slot to slot, the number of packets awaiting transmission (called backlog) is a Markov chain taking values in $\{0, 1, 2, \quad\}$. The central problem is to investigate under what conditions the backlog Markov chain is *ergodic*, i.e., it is stable in the sense that it reaches a steady state in which the periods between the times when there are no packets to transmit are not too infrequent (they have finite expected value). The transition probabilities of the Markov chain are parametrized by the rate of arrival of new packets λ and the retransmission probability p. Whereas λ is fixed and given, p is chosen by the transmitters. Hence, we are dealing with a fairly simple *controlled Markov chain* whose control space is the interval $(0, 1]$. In the original ALOHA algorithm, the control p remained constant and common to all transmitters regardless of the information acquired by listening to the channel, thereby resulting in the open-loop control of the Markov chain. Depite several "proofs" of the stability of ALOHA published during the 1970's, neither the actual system built in Hawaii nor the ideal Markov chain model were stable. The reason why the open loop system is unstable can be easily understood by considering the backlog drift, $d(n)$, which is defined as the expected increase in the backlog over the next slot when the current value of the backlog is equal to n. It is easy to see that the backlog drift is given simply by the expected number of new packets per slot minus the expected number of successfully transmitted packets in the next slot, i.e.,

$$d(n)=\lambda-\{np(1-p)^{n-1}\}. \tag{3.1}$$

The drift quantifies the expected evolution of the Markov chain from each state, and therefore it is a valuable tool in analyzing the stability of the chain. For any $p\in(0, 1]$ the term in brackets in

(3.1) goes to 0 as $n \to \infty$, and hence, the drift is positive and close to λ for sufficiently large backlogs. This implies that when the backlog is large it tends to grow, thereby eliminating any hope for stability. Using standard results, this reasoning can be formalized straightforwardly to prove not only the instability of the open-loop system [11] for all values of λ and p, but the fact that the backlog goes to infinity with probability one [25], [35], [40].

Fortunately, the system can be stabilized by closed-loop control. Let us examine first the case of complete-state information, i.e., each station is informed at the end of each slot of the current value of the backlog and chooses the retransmission probability on the basis of that information. As far as stability is concerned, the best choice of the retransmission probability p is the value that minimizes the drift because that results in the maximum possible arrival rate that guarantees stability (called the throughput). It follows from (3.1) that the optimum value of p is

$$p^*(n) = \frac{1}{n}, \qquad n = 1, 2, \cdots \tag{3.2}$$

and the resulting drift is

$$d^*(n) = \lambda - \left[1 - \frac{1}{n}\right]^{n-1} \tag{3.3}$$

which is negative for $n > 1$ when $\lambda < e^{-1}$, and is positive for large backlogs when $\lambda > e^{-1}$. Therefore, the throughput of the closed-loop system with complete state information is $e^{-1} = 0.368$. However, the relevance of complete state information feedback is rather limited in practice. This is because the instantaneous value of the backlog is available to each station only if there exists so large a degree of communication among the transmitters that much more efficient algorithms than ALOHA can be used.

The case of partial state information is the problem of interest in practice, since the only feedback available to each station is the outcome (collision, success, empty) of the transmission in each slot. The analysis of the controlled system with partial state information was pioneered by Hajek and Van Loon [20] who proposed a recursive updating law of the retransmission probabilities as a function of the channel outcomes. This feedback policy was shown in [21] to attain the throughput achievable with complete-state information, namely e^{-1}. Those papers and subsequent works have referred to the problem as *decentralized control* of ALOHA, motivated by the fact that each station chooses the retransmission probability autonomously based on the channel feedback. However, it is useful to recognize that the problem boils down to (centralized) stochastic control with one decision maker and incomplete state information because all stations are constrained to use the same retransmission probabilities.

We will review here the proof of stability of the following *certainty-equivalence* closed-loop control:

$$p(\hat{n}) = \frac{1}{\hat{n}} \tag{3.4}$$

where $\hat{n}$ is an estimate of the backlog updated according to

$$\hat{n}_{k+1} = \begin{cases} \max\{1, \hat{n}_k + \alpha\} & k\text{th slot is idle} \\ \hat{n}_k + \beta & k\text{th slot is success or collision.} \end{cases} \tag{3.5}$$

The throughput attainable with this feedback law depends on the constants $\alpha < 0$ and $\beta > 0$. As we will see, there exists a set of choices for those constants that results in throughput equal to e^{-1}.

Unlike the case of complete-state information, the proof of stability is not straightforward because now it is the two-dimensional process formed by the backlog and its estimate $\{(n_k, \hat{n}_k)\}_k$ (rather than the backlog itself) which is a Markov process. According to (3.4) and (3.5) the drift of this Markov process is given by

$$E[(n_{k+1}, \hat{n}_{k+1}) - (n_k, \hat{n}_k) | (n_k, \hat{n}_k) = (n, s)]$$
$$= \left(\lambda - \frac{n}{s}\left[1 - \frac{1}{s}\right]^{n-1}, \beta + (\max\{\alpha, 1-s\} - \beta)\left[1 - \frac{1}{s}\right]^{n}\right)$$
$$\triangleq (d(n, s), c(n, s)). \tag{3.6}$$

Contrary to what we saw in the case when the state is known, it is not true that the backlog drift is negative for sufficiently large backlogs. As we can see in Fig. 5, if the estimate is far from the true value, then the backlog may actually tend to increase.

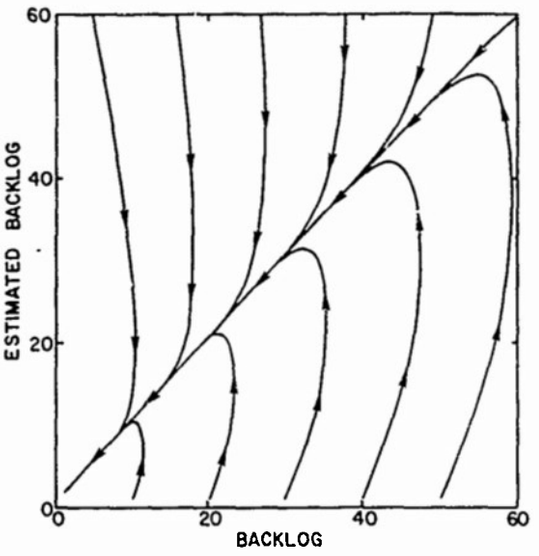

Fig. 5. Drift of (backlog, backlog estimate) Markov process for decentralized control with $\alpha = -1.48$, $\beta = 0.8$, and $\lambda = 0.33$.

However, at every point in the state space the tendency of the process is to approach the diagonal where the estimate is equal to the true value of the backlog. Furthermore, as Fig. 5 or the analysis of the perfect-state information case shows, the drift along the diagonal is negative. Such a behavior is a strong indication of the stability of the controlled Markov process.

This can be proved using a powerful sufficient condition found by Mikhailov [31] for the stability of a Markov process taking values in $\mathbb{R}^+ \times \mathbb{R}^+$. In essence, Mikhailov's condition states that it is enough to restrict attention to those points of the state space where either the backlog or its estimate are large and at which the drift is *radial*, i.e.,

$$\frac{d(n, s)}{c(n, s)} = \frac{n}{s};$$

then, it is sufficient for stability that the drift point towards the origin at those states. To see that this condition is indeed satisfied for our system, we compute first the asymptotic drifts along the radius $\{(n, s): n/s = \psi\}$ for $\psi \in [0, \infty)$

$$d(\psi) = \lim_{s \to \infty} d(\psi s, s) = \lambda - \psi e^{-\psi} \tag{3.7a}$$

$$c(\psi) = \lim_{s \to \infty} c(\psi s, s) = \beta + (\alpha - \beta)e^{-\psi}. \tag{3.7b}$$

It can be checked using (3.7) that if the constants α and β in (3.5) are chosen such that $\beta > 0.23\lambda$ and $\lambda - e^{-1} = \beta + (\alpha - \beta)e^{-1}$, then the drift is radial only at $\psi = 1$ (cf. Fig. 5), where it points towards the origin as long as $d(1) = \lambda - e^{-1} < 0$.

Mikhailov's sufficient condition can be justified constructing a

stochastic Lyapunov function to prove the stability of a Markov process $\{x_k\}_k$ with state space $\mathbb{R}^+ \times \mathbb{R}^+$. To that end, it is advantageous to switch to polar coordinates (r, ϕ) and to define the radial drift $\delta(r, \phi)$ as the projection of the drift along the direction of the point (r, ϕ) and the tangential drift $\mu(r, \phi)$ as the projection of the drift along the direction perpendicular to (r, ϕ). Denote the asymptotic drifts $\delta(\phi) = \lim_{r \to \infty} \delta(r, \phi)$ and $\mu(\phi) = \lim_{r \to \infty} \mu(r, \phi)$ and define the function

$$V(r, \phi) = r\Phi(\phi)$$

where

$$\Phi(\phi) = \exp\left(-C \int_0^{\phi} \mu(v)\, dv\right) \qquad \phi \in \left[0, \frac{\pi}{2}\right].$$

Note that $V(r, \phi)$ is a candidate Lyapunov function because it is positive outside the origin and $V(r, \phi) \to \infty$ as $r \to \infty$. Furthermore, it can be shown [31] that the asymptotic drift of the candidate Lyapunov function is equal to

$$\lim_{r \to \infty} E[V(x_{k+1}) - V(x_k) | x_k = (r, \phi)] = \Phi(\phi)[\delta(\phi) - C\mu^2(\phi)]. \quad (3.8)$$

Now, under Mikhailov's condition, the asymptotic drifts are assumed continuous on $[0, \pi/2]$ and $\delta(\phi) < \epsilon$ for any phase such that $\mu(\phi) = 0$ (i.e., whenever the drift is radial it points towards the origin), therefore, the constant C can be chosen large enough so that the left side of (3.8) is upper bounded by a negative constant. This implies that $V(r, \phi)$ is indeed a stochastic Lyapunov function and therefore standard results on the stability of stochastic systems [27], [45] can be applied to show the stability of the system.[4]

In some multiaccess environments, the receiver can indeed demodulate reliably one or more packets even in the presence of other interfering packets and the collision channel model no longer applies to those cases. The results reviewed in this section can be generalized to a general channel with *multipacket reception* capability, to show that: 1) the throughput of open-loop ALOHA is equal to the limit of the expected number of successfully received packets per slot as the backlog goes to infinity [17]; and 2) the throughput of closed-loop ALOHA (with either complete or partial state information) is equal to the maximum over v of the expected number of successfully received packets per slot when the number of attempted transmissions is a Poisson random variable with mean v [18].

Returning to the case of the collision channel, the next natural step is to drop the main restriction in the ALOHA algorithm, namely, that all stations use the same retransmission probability. This is done in a class of random-access algorithms referred to as collision resolution algorithms which are characterized by the fact that not only are all blocked packets eventually retransmitted successfully, but all users eventually become aware that these packets have been successfully retransmitted. Contrary to the ALOHA algorithm, the decision whether or not to transmit a packet takes into account the previous history of attempted retransmissions of that particular packet. The introduction of this new dimension into the problem renders Markov chain tools considerably less useful than in the foregoing analysis and converts it into a very difficult decentralized stochastic control problem, for which the optimum throughput remains unknown[5] despite many efforts.

[4] Another choice of stochastic Lyapunov function for the specific case of decentralized control of ALOHA can be found in [43].

[5] The best known algorithm has been shown to achieve a throughput of 0.488 using Howard's policy iteration for sequential infinite-horizon problems [32] or by reduction to a simple optimization problem [48]. On the other hand, it is known that the optimum throughput is upper bounded by 0.568 [44].

B. Simultaneous Transmission

In contrast to random-access communication systems, in simultaneous transmission multiple-access systems, the transmitters send their messages simultaneously, independently, and without monitoring the channel in any way. The most common type of simultaneous transmission system is code-division multiplexing, where each user modulates a preassigned signature waveform known by the receiver.

Specifically, we will assume that in order to send the message $\{b_k(i) \in A\}_{i=0}^{M-1}$ (i.e., a string of M symbols drawn from a finite set A), the kth user transmits

$$\sum_{i=0}^{M-1} b_k(i) s_k(t - iT)$$

where $\{s_k(t), 0 \le t \le T\}$ is the waveform assigned to the kth user, and T is the symbol period. Then the demodulator receives the sum of the signals transmitted by the K active users embedded in noise

$$r(t) = \sum_{k=1}^{K} \sum_{i=0}^{M-1} b_k(i) s_k(t - iT - \tau_k) + n(t) \quad (3.8)$$

where the offsets $\tau_{k-1} \le \tau_k \in [0, T)$ model the fact that the users do not synchronize their transmissions. Then the task of the receiver is to recover the transmitted information strings $\{b_k(i)\}_{i=0}^{M-1}{}_{k=1}^{K}$. Following [47] we will show how to obtain an optimum multiuser demodulator via dynamic programming. First, denote the MK-vector

$$d = \{d_{k+iK} = b_k(i),\ k = 1, \cdots, K,\ i = 0, \cdots, M-1\}$$

and the multiuser signal in (3.8)

$$S(t, d) = \sum_{k=1}^{K} \sum_{i=0}^{M-1} b_k(i) s_k(t - iT - \tau_k) = \sum_{i=1}^{MK} d_i z_i(t) \quad (3.9)$$

where $z_{k+iK}(t) = s_k(t - iT - \tau_k)$.

A reasonable criterion for demodulating the information carried in $S(t, d)$ upon observation of $r(t)$ is to select the MK-vector d that best explains the received waveform in the sense of minimizing the energy of the corresponding noise realization, i.e.,

$$\min_{d \in A^{MK}} \|S(t, d) - r(t)\|^2. \quad (3.10)$$

If the noise $n(t)$ is white and Gaussian, then this criterion results in maximum likelihood decisions. Equivalently, the objective is to find the vector that solves

$$\max_{d \in A^{MK}} \Omega(d) \quad (3.11)$$

where

$$\Omega(d) = 2 \int_{-\infty}^{\infty} S(t, d) r(t)\, dt - \int_{-\infty}^{\infty} S^2(t, d)\, dt. \quad (3.12)$$

Since the maximization in (3.11) is over a finite set, we could solve it by the brute-force method of evaluating $\Omega(d)$ for each possible argument. However, it is possible to decompose $\Omega(d)$ in a sequential fashion that lends itself to efficient optimization. From

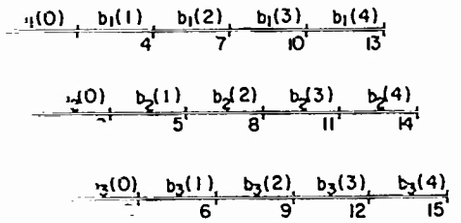

Fig. 6. Symbol epochs for $K = 3$ and $M = 5$.

3.9) it is immediate to write the first integral in (3.12) sequentially

$$\int_{-\infty}^{\infty} S(t, d)r(t)\,dt = \sum_{j=1}^{MK} d_j y_j \tag{3.13}$$

where

$$y_j = \int_{-\infty}^{\infty} z_j(t)r(t)\,dt. \tag{3.14}$$

This implies that the objective function (3.12) depends on $r(t)$ only through the quantities $\{y_j\}_{j=1}^{MK}$, which are obtained by correlating $r(t)$ with each of the signature waveforms during each symbol epoch. In order to find an explicit expression for the second integral on the right-hand side of (3.12), which is the energy of the multiuser signal, we will denote

$$R(j, l) = \int_{-\infty}^{\infty} z_j(t)z_l(t)\,dt. \tag{3.15}$$

It follows immediately from the definition that these coefficients satisfy the following properties.

1) $R(k + iK, k + iK) = \int_0^T s_k^2(t) \triangleq w_k$.
2) $R(k + iK, n + iK) = R(k, n)$ for all i.
3) $R(j, l) = 0$ unless $|j - l| < K$.

The first property indicates that each of the diagonal elements of $R(i, j)$ is equal to the energy of one of the K assigned waveforms. The second and third properties can be illustrated by referring to Fig. 6 which represents the symbol epochs of three asynchronous users sending strings of $M = 5$ symbols. Each symbol period in Fig. 6 is labeled with the index of the corresponding component of the vector d. The second property indicates that the cross-correlations between two signals depend only on their relative location (e.g., $R(4, 6) = R(13, 15)$ in Fig. 6) and the third property states that each symbol only interferes with $2K - 2$ symbols of the other users [e.g., in Fig. 6, $d_9 = b_3(2)$ only overlaps with $d_7 = b_1(2)$, $d_8 = b_2(2)$, $d_{10} = b_1(3)$, and $d_{11} = b_2(3)$]. It follows from these properties that the coefficients in (3.15) can be obtained from the $K \times K$ matrix $\{R(k, n)\}_{k=1, n=1}^{K, K}$ whose diagonal elements correspond to the energy per symbol of each user and whose off-diagonal elements correspond to the cross-correlations between the signature waveforms of each pair of users. Using (3.15), the foregoing properties, and letting $\kappa(j) \in \{1, \cdots, K\}$ be the modulo-K remainder of j (i.e., for some i, $j = \kappa(j) + iK$), we can write

$$\begin{aligned}\int_{-\infty}^{\infty} S^2(t, d)\,dt &= \sum_{j=1}^{MK}\sum_{l=1}^{MK} d_j d_l R(j, l) \\ &= \sum_{j=1}^{MK} d_j \left[w_{\kappa(j)} + 2\sum_{l=j-K+1}^{j-1} d_l R(j, l) \right] \\ &= \sum_{j=1}^{MK} d_j \left[w_{\kappa(j)} + 2\sum_{n=1}^{K-1} d_{j-n} g_{\kappa(j)}(K-n) \right]\end{aligned} \tag{3.16}$$

where $g_k(m) = R(k + K, k + m)$. Putting together (3.12), (3.13), and (3.16) we see that we can express $\Omega(d)$ as a sum of MK terms, each of which depends on K components of d and such that consecutive terms depend on the same components but one. Specifically, we can write

$$\Omega(d) = \sum_{j=1}^{MK} \lambda_j(x_j, d_j) \tag{3.17}$$

where

$$\lambda_j(x, u) = u[2y_j + uw_{\kappa(j)} - 2x^T g_{\kappa(j)}] \tag{3.18}$$

and x_j is the state of a shift-register $K - 1$ dimensional system

$$x_{j+1}^T = [x_{j+1}(1), \cdots, x_{j+1}(K-1)] = [x_j(2), \cdots, x_j(K-1), d_j];$$
$$x_0 = 0. \tag{3.19}$$

It is now apparent that the solution to (3.11) entails solving a *finite-horizon deterministic optimal control problem* with additive costs per stage for the linear system in (3.19), and with a finite admissible control set A. Therefore, optimum multiuser demodulation is equivalent to a shortest path problem in an M-stage layered directed graph, where at each stage there are A^{K-1} states. This optimization problem can be solved by dynamic programming (e.g., [7]) in backward or forward fashion. In practice, it is necessary to demodulate the transmitted symbols in real-time, and since M is usually a very large integer, it is not feasible to wait until all the observables $\{y_j\}_{j=1}^{MK}$ have been obtained before starting to make decisions. Therefore, a suboptimum version of the forward dynamic programming algorithm is adopted in practice whereby each decision is based on the paths corresponding to the cost-to-arrive function computed a fixed number of steps ahead. This real-time version of forward dynamic programming is known in communication theory as the Viterbi algorithm [12], and was originally devised (without resorting to the dynamic programming framework) for decoding convolutional codes. The maximum-likelihood criterion used in (3.10) is not the only possible optimality criterion. For example, if the objective is to minimize the probability of error for each user, then the multiuser demodulator uses a *backward-forward* dynamic programming algorithm [49] whereby optimum decisions are based on the independent computation of a cost to go and a cost-to-arrive function.

IV. Other Problem Areas

Routing and multiple access are not the only problem areas in the field of communication networks which control theory can help formulate, study, and solve. We have deliberately chosen to confine our attention to these two areas in order to get across in a concise manner our belief that the field of communication networks offers a rich selection of applications for control theory. We would feel remiss, however, if we did not even make an attempt to provide a taste of some of the numerous other design and operation issues that, again, bring forth control systems concepts and techniques. For this purpose, and with a conscious effect not to expand in depth but only to describe, we will mention two areas from point-to-point networks and one from radio networks. The first two concern flow control and integrated switching, respectively, while the third concerns the problem of scheduling transmission in multihop networks. Unlike the cases of routing and multiple access, these areas have not yet fully benefitted from the use of control theoretic approaches although such approaches would be very well suited to them indeed.

A. Flow Control

A stark reality in the design of networks is that despite the reduction of the cost of memory, storage at each node is going to be finite. Coupled with another reality, namely that data transmissions on the whole continue to be bursty, it implies that buffer overflow may occur and, along with it, congestion and deadlocks. Flow control is the name we use to describe the collection of measures taken to avoid buffer overflow and highly congested nodes in the network. Congestion and saturation are often the consequences of diverging, unstable behavior. Thus, it is of interest not only to optimize over possible flow control strategies, but to determine their robustness against disturbances or modeling inaccuracies that may lead to unstable behavior.

The control variables in flow control problems are admission (or blocking) probabilities for messages or sessions at the source node. In practice these are often implemented in terms of a bang-bang control strategy known as *window flow control* whereby input ports are allowed to continuously inject messages into the network at the full desired input rate until the number of *unacknowledged*[6] messages exceeds the value of the "window size" w. A simple, yet unanswered question is, what should the value of w be?

Previous efforts to use control theory tools to analyze optimal flow control problems include [28] and [46] where the optimality of window flow control is proved within the domain of a simplified model, and [39] where dynamic programming value iteration techniques are used to characterize optimal flow control performance. An alternative approach to the flow control problem is to subsume it into the static routing problem considered in Section II-A [19]: suppose that for every source-destination pair a fictitious direct link is added between them. We can then interpret the blocking action of a flow control procedure as a diversion of the blocked portion of the traffic through this fictitious link to the destination. Thus, we can consider that no traffic is blocked. Of course, in order to discourage the use of this fictitious link we must augment the overall delay cost function with a term that penalizes appropriately the use of this link.

B. Integrated Switching

A revolutionary development in the field of networks whose implementation is currently under way is the combination of the capabilities of what have been separately developed in the past and called voice networks and data networks. Voice is a commodity that must meet different requirements than data. For example, speech signals have inherent redundancy that make them quite robust with respect to occasional errors or deliberate compression. At the same time, except in applications of voice messaging, speech signals occur in the context of real-time conversations and, as such, must encounter short and, more importantly, constant delay. On the other hand, data must preserve their integrity and cannot tolerate errors; however, long and variable delays can be often tolerated.

How does one design a single network that can handle such dissimilar commodities with automated procedures? The natural course of events in the last decade or two was to attempt to force data on primarily voice networks or to let voice ride on what were mainly data networks. The literature is full of ideas for baseline integration that are mostly heuristic and difficult to analyze. An attempt to formulate the problem of integrated switching as an optimization problem was presented in [50]. In its simplest form the model is as follows: consider a single node in the network with a single outgoing link on which incoming voice calls and data packets must be multiplexed. Let W be the bandwidth of the outgoing link. Let V be the bandwidth required for the continuous, uninterrupted accommodation of a single voice call. Let, therefore, $N = W/V$ be the maximum number of calls that can be assigned dedicated circuits simultaneously if no data packets are transmitted. A voice call can either be accepted (and assigned the necessary bandwidth V) or blocked. Data packets can be stored in a buffer facility. If, at a given time, there are i calls in the system, the data packets can be served at the full rate corresponding to the remaining bandwidth $W - iV$. Such a switching architecture represents what has been called the *movable boundary* idea in integration. A natural MDP can be simply formulated as follows: choose the control action of blocking or accepting a call upon arrival in order to minimize the weighted sum of the average data packet delay and the call-blocking probability. If we assume that both arrival streams (voice calls and data) are independent Poisson processes, that the call holding time is exponentially distributed, and that the message lengths are likewise exponential, we can apply the technique described in Section II of converting the MDP to an LP and show that the optimal policy has the useful *switching-type* form. Namely, if i is the number of ongoing calls and j the total number of data messages at the node, the optimal control action should be to block the call in region B of the state space as shown in Fig. 7 and to accept it in region A.

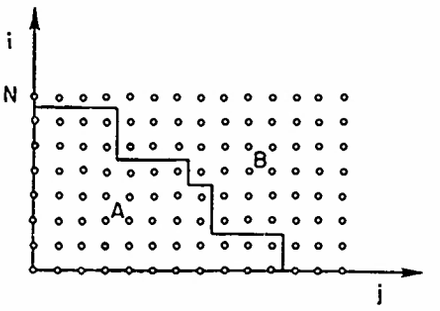

Fig. 7. Switching-type optimum policy for integrated switching.

C. Link Scheduling

Let us now turn our attention back to the radio network environment. In Section III the multiple access channel was considered and a number of difficult but interesting control problems were identified. Throughout that discussion, it was assumed that all terminals are within a single transmission hop from the destination. In many radio networks, however, this is not the case. Messages need to be relayed via intermediate nodes to their final destinations. Thus, the familiar problem of routing arises again, except that this time there is a new twist to it. In point-to-point networks, transmissions between different node pairs can take place simultaneously because there are dedicated, "hard-wired" links between the corresponding nodes. In a radio (or, more generally, in a multiaccess/broadcast) environment, if the nodes are densely connected, not all transmissions can take place simultaneously (unless separate dedicated channels or simultaneous transmission signaling techniques (Section III-B) are used). They must be scheduled in time to avoid the interference that would occur otherwise.

It becomes evident that the mere fact that the transmission among a group of nodes must take place one at a time raises the question whether the intended transmissions are routing-wise optimal any more. Several versions of this problem have been studied [3], [23], [36]. In every case and even if the routing problem is sidestepped, we are led to hard combinatorial optimization problems where questions of computational complexity and distributed implementation are of primary importance.

V. Conclusion

It should be clear by now that the theory of linear and nonlinear optimization, dynamic programming, stochastic control, stability

[6] Note the implicit assumption of delayed feedback information from the destination to the source node.

analysis, and distributed control have found interesting applications arising in the analysis and design of communication networks. Unlike other complex systems that have been successfully studied by control system theorists in the past (such as chemical plants, flexible aircraft, robot systems, etc.), communication networks stand out in that the commodity to be controlled is information (including its transmission, storage, processing, etc.). This feature, perhaps, misleads and intimidates those who do not feel sufficiently inter-disciplinarian to tackle these problems. We hope that by having selected to present a few examples in which concrete, purely control-theoretic problems can be formulated and have been (or can be) studied successfully, we may encourage attention by the control community to this application area that is especially rich in new challenges.

As stated from the outset, we did not attempt to survey or completely cover the multiple control facets of communication networks. The collection in this paper merely represents an effort to illuminate a few selected problem areas and to show how control techniques apply to them.

References

[1] N. Abramson, "Development of the ALOHANET," *IEEE Trans. Inform. Theory*, vol. IT-31, pp. 119-123, Mar. 1985.

[2] P. O. Bhattacharya *et al.* "A (not very) simple dynamic routing problem," in *Proc. 25th Allerton Conf. Contr. Commun. Comput.*, Urbana, IL, 1987, pp. 998-1006.

[3] D. J. Baker, J. Wieselthier, and A. Ephremides, "A distributed algorithm for scheduling the activation of links in a self-organizing mobile radio network," in *Proc. IEEE Int. Conf. Commun.*, Philadelphia, PA, June 1982, pp. 2F.6.1-5.

[4] M. Bazaraa and C. Shetty, *Nonlinear Programming: Theory and Algorithms*. New York: Wiley, 1979.

[5] D. Bertsekas and R. Gallager, *Data Networks*. Englewood Cliffs, NJ: Prentice-Hall, 1987.

[6] D. P. Bertsekas, "Distributed dynamic programming," *IEEE Trans. Automat. Contr.*, vol. AC-27, pp. 610-616, June 1982.

[7] D. P. Bertsekas, *Dynamic Programming: Deterministic and Stochastic Models*. Englewood Cliffs, NJ: Prentice-Hall, 1987.

[8] C. Buyukkoc, P. Varaiya, and J. Walrand, "The μc rule revisited," *Advances Appl. Probability*, vol. 17, pp. 237-238, 1985.

[9] C. A. Courcoubetis, "Optimal control of a queueing system with simultaneous service requirements," *IEEE Trans. Automat. Contr.*, vol. AC-32, pp. 717-727, 1987.

[10] A. Ephremides, P. Varaiya, and J. Walrand, "A simple dynamic routing problem," *IEEE Trans. Automat. Contr.*, vol. AC-25, pp. 690-693, Aug. 1980.

[11] G. Fayolle, E. Gelenbe, and L. Labetoulle, "Stability and optimal control of the packet switching broadcast channel," *JACM*, vol. 24, pp. 375-386, 1977.

[12] G. D. Forney, "The Viterbi algorithm," *Proc. IEEE*, vol. 61, pp. 268-278, Mar. 1973.

[13] M. Frank and P. Wolfe, "An algorithm for quadratic programming," *Naval Res. Logist. Quat.*, vol. 3, pp. 149-154, 1956.

[14] L. Fratta, M. Gerla, and L. Kleinrock, "The flow deviation method: An approach to store-and-forward communication network design," *Networks*, vol. 3, pp. 97-133, 1973.

[15] T. Gal, *Postoptimal Analyses, Parametric Programming, and Related Topics*. Englewood Cliffs, NJ: McGraw-Hill, 1979.

[16] R. Gallager, "A minimum delay routing algorithm using distributed computation," *IEEE Trans. Commun.*, vol. COM-23, pp. 73-85, 1977.

[17] S. Ghez, S. Verdu, and S. Schwartz, "Stability properties of slotted Aloha with multipacket reception capability," *IEEE Trans. Automat. Contr.*, vol. 33, pp. 640-649, July 1988.

[18] S. Ghez, S. Verdu, and S. Schwartz, "Optimal decentralized control in the multipacket channel," *IEEE Trans. Automat. Contr.*, to be published.

[19] S. J Golestaani, "A unified theory of flow control and routing on data communication networks," Ph.D. dissertation, Mass. Inst. Technol., Cambridge, 1980.

[20] B. Hajek and T van Loon, "Decentralized dynamic control of a multiaccess broadcast channel," *IEEE Trans. Automat. Contr.*, vol. AC-27, pp. 559-569, June 1982.

[21] B Hajek, "Hitting-time and occupation-time bounds implied by drift analysis with applications," *Adv. Appl. Prob.*, vol. 14, pp. 502-525, 1982.

[22] ——, "Optimal control of two interacting service stations," *IEEE Trans. Automat. Contr.*, vol AC-29, pp 491-499, June 1984.

[23] B Hajek and G Sasaki, "Link scheduling in polynomial time," *IEEE Trans. Inform. Theory*, vol. 34, pp. 910-917, Sept 1988.

[24] J. M. Harrison, "Dynamic scheduling of a multiclass queue: Discount optimality," *Oper. Res.*, vol. 23, pp. 270-282, 1975.

[25] F P Kelly, "Stochastic models of computer communications systems," *J. R. Statist. Soc. B*, vol. 47, pp. 379-395, 1985.

[26] G. P. Klimov, "Time sharing service systems," *Theory Probability Appl.*, vol. 19, pp. 532-551, Sept. 1974; see also vol. 23, pp. 314-321, June 1978.

[27] H Kushner, *Introduction to Stochastic Control*. New York: Holt, Rinehart, and Winston, 1971.

[28] A. Lazar, "Optimum flow control of a class of queueing networks in equilibrium," *IEEE Trans. Automat. Contr.*, vol. AC-28, pp. 1001-1007, Nov. 1983.

[29] W Lin and P R Kumar, "Optimal control of a queueing system with two heterogeneous servers," *IEEE Trans. Automat. Contr.*, vol. AC-29, pp. 696-703, Aug. 1984.

[30] S. A. Lippman, "Semi-Markov decision processes with unbounded rewards," *Management Sci.*, vol. 19, pp. 717-731, 1973.

[31] V A. Mikhailov, "Geometrical analysis of the stability of Markov chains in R^n and its application to throughput evaluation of the adaptive random multiple access algorithm," *Problems of Inform Transmission*, vol. 24, pp. 47-56, Jan.-Mar. 1988.

[32] J. Mosely and P. Humblet, "A class of efficient contention resolution algorithms for multiple access channels," *IEEE Trans. Commun.*, vol. COM-33, pp.145-151, Feb. 1985.

[33] P. Nain and K. W. Ross, "Optimal priority assignment with hard constraint," *IEEE Trans. Automat. Contr.*, vol. AC-31, pp. 883-888, 1986.

[34] C. H. Papadimitrou and K Steiglitz, *Combinatorial Optimization. Algorithms and Complexity*. Englewood Cliffs, NJ: Prentice-Hall, 1982.

[35] S. Parekh, F. Schoute, and J. Walrand, "Instability and geometric transience of the Aloha protocol," in *Proc. 26th Conf. Decision Contr.*, Los Angeles, CA, Dec. 1987, pp. 1073-1077.

[36] M. J. Post, P. E. Sarachik, and A. S. Kershenbaum, "A biased greedy algorithm for scheduling multi-hop radio networks," in *Proc. Conf. Inform. Sci. Syst.*, Johns Hopkins Univ., Baltimore, MD, 1985, pp 564-672.

[37] R. T. Rockafellar, *Convex Analysis*. Princeton, NJ: Princeton University Press, 1970.

[38] Z. Rosberg, P. Varaiya, and J. Walrand, "Optimal control of service in tandem queues," *IEEE Trans. Automat. Contr.*, vol. AC-27, pp. 600-610, June 1982.

[39] Z. Rosberg and I. Gopal, "Optimal hop-by-hop flow control in computer networks," *IEEE Trans. Automat. Contr.*, vol AC-31, pp. 813-822, Sept. 1986.

[40] W. Rosberg and D Towsley, "On the instability of slotted-ALOHA multiaccess algorithm," *IEEE Trans. Automat. Contr.*, vol AC-28, pp. 994-996, Oct. 1983.

[41] R. E. Tarjan, *Data Structures and Network Algorithms* (CBMS-NSF Reg. Conf. Series in Appl. Math. no. 7). Philadelphia, PA: SIAM, 1983.

[42] J. N. Tsitsiklis and D. P. Bertsekas, "Distributed asynchronous optimal routing in data networks," *IEEE Trans. Automat. Contr.*, vol. AC-31, pp. 325-332, Apr. 1986.

[43] J. N. Tsitsiklis, "Analysis of a multiaccess control scheme," *IEEE Trans. Automat. Contr.*, vol. AC-32, pp. 1017-1020, Nov 1987

[44] B. S. Tsybakov and N. B. Likhanov, "An improved upper bound on capacity of the random multiple-access channel," *Problemi Pederachi Informatsii*, vol. 23, pp. 64-78, 1987.

[45] R. Tweedie, "Sufficient conditions for ergodicity and recurrence of Markov chains on a general state space," *Stoch. Proc. Appl.*, vol. 3, pp. 385-403, 1975.

[46] F. Vakil and A. Lazar, "Flow control protocols for integrated networks with partially observed voice traffic," *IEEE Trans. Automat. Contr.*, vol. AC-32, pp. 2-14, Jan. 1987.

[47] S. Verdu, "Minimum probability of error for asynchronous Gaussian multiple-access channels," *IEEE Trans. Inform. Theory*, vol. IT-32, pp. 85-96, Jan. 1986.

[48] ——, "Computation of the efficiency of the Mosley-Humblet contention resolution algorithm. A simple method," *Proc. IEEE*, vol 74, pp. 613-614, Apr. 1986.

[49] S Verdu and H. V. Poor, "Abstract dynamic programming models under commutativity conditions," *SIAM J. Contr. Optimiz.*, vol. 4, July 1987.

[50] I. Viniotis and A. Ephremides, "Optimal switching of voice and data at a network node," *Proc. 26th CDC*, Los Angeles, CA, Dec. 1987, pp. 1504-1507.

[51] J. Viniotis, Ph.D. dissertation, Univ. Maryland, College Park, MD, 1988.

[52] J. Walrand, "A note on optimal control of a queueing system with two heterogeneous servers," *Syst. Contr. Lett.*, vol. 4, pp. 131-134, 1984.

[53] ——, *An Introduction to Queueing Networks*. Englewood Cliffs, NJ: Prentice-Hall, 1988.

[54] R. Weber, "On the optimal assignment of customers to parallel servers," *J. Appl. Prob.*, vol. 15, pp. 406-413, 1978.

[55] Z. Wu, P. B. Luh, S. Chang, and D. A. Castanon, "Optimal control of

queueing system with two interacting service stations and three classes of impatient tasks," *IEEE Trans. Automat. Contr.*, vol. 33, pp. 42–49, 1988.

Anthony Ephremides (S'68–M'71–SM'77–F'84) was born in Athens, Greece, in 1943. He received the Ph.D. degree in electrical engineering from Princeton University, Princeton, NJ, in 1971.

He has been with the University of Maryland, College Park, since 1971. He has also spent semesters on leave at M.I.T., the University of California, Berkeley, and ETH, Zurich. He is active in professional consulting as President of Pontos, Inc. Currently, his research interests lie in the areas of communication systems, performance analysis, modeling, optimization, simulation, and design.

Dr. Ephremides is Director of Division X of the IEEE and has served as President of the Information Theory Society and on the Board of Governors of the Control Systems Society. He has been Associate Editor of the IEEE TRANSACTIONS ON AUTOMATIC CONTROL and General Chairman of major IEEE Conferences.

Sergio Verdú (S'80–M'84–SM'88) was born in Barcelona, Catalonia, Spain, in 1958. He received the Telecommunication Eng. degree from the Polytechnic University of Barcelona in 1980 and the Ph.D. degree in electrical engineering from the University of Illinois at Urbana-Champaign in 1984.

Upon completion of his doctorate he joined the faculty of Princeton University, Princeton, NJ, where he is currently an Associate Professor of Electrical Engineering. His current research interests are in the areas of multiuser communication and information theory.

Dr. Verdú is a recipient of the National University Prize of Spain, the Rheinstein Outstanding Junior Faculty Award of the School of Engineering and Applied Science at Princeton University, and the NSF Presidential Young Investigator Award. He is currently serving as Associate Editor of the IEEE TRANSACTIONS ON AUTOMATIC CONTROL, and as a member of the Board of Governors of the IEEE Information Theory Society.

T-AC/34/11//30788

Optimal Decentralized Control in the Random Access Multipacket Channel

Sylvie Ghez
Sergio Verdú
Stuart C. Schwartz

Reprinted from
IEEE TRANSACTIONS ON AUTOMATIC CONTROL
Vol. 34, No. 11, November 1989

Optimal Decentralized Control in the Random Access Multipacket Channel

SYLVIE GHEZ, STUDENT MEMBER, IEEE, SERGIO VERDÚ, SENIOR MEMBER, IEEE, AND STUART C. SCHWARTZ, SENIOR MEMBER, IEEE

Abstract—A decentralized control algorithm is sought that maximizes the stability region of the infinite-user slotted multipacket channel and is easily implementable. To this end, the perfect state information case where the stations can use the instantaneous value of the backlog to compute the retransmission probability is studied first. The best throughput possible for a decentralized control protocol is obtained, as well as an algorithm that achieves it. Those results are then applied to derive a control scheme when the backlog is unknown, which is the case of practical relevance. This scheme, based on a binary feedback, is shown to be optimal given some restrictions on the channel multipacket reception capability.

I. INTRODUCTION

MOST studies on random access communications rely on the assumption that when two or more packets overlap, all the information that was sent is irremediably lost, hence the need to repeat all transmissions at some later time. This is actually a pessimistic point of view, since there are many examples of random access systems where one or more packets may be successful in the presence of other simultaneous transmissions. In order to represent such random access systems, a model for a channel with multipacket reception capability has been developed in [6]-[8]. We consider a slotted channel with an infinite population of users, and we assume that the probability of having k successes in a slot where there are n transmissions depends only on the collision size n

$$\epsilon_{nk} = P[k \text{ packets are correctly received} | n \text{ are transmitted}] \qquad (n \geq 1,\ 0 \leq k \leq n).$$

We define the reception matrix as

$$E = \begin{bmatrix} \epsilon_{10} & \epsilon_{11} & & & \\ \epsilon_{20} & \epsilon_{21} & \epsilon_{22} & 0 & \\ \cdot & \cdot & & & \\ \epsilon_{n0} & \epsilon_{n1} & & \epsilon_{nn} & \\ \cdot & \cdot & & & \cdot \\ \cdot & \cdot & & & \cdot \end{bmatrix}.$$

This model can be applied to channels with capture [1]-[3], [10], [16], [18], [20], [23], [26], [28], [34] and to systems using CDMA [22], [24], [29]. It is also relevant for many other applications, such as systems with multiuser detectors [33] or, for instance, the channel studied in [17], [31]. For more details about this model, the reader is referred to [6] and [8]. Denoting by $C_n = \sum_{k=1}^{n} k\epsilon_{nk}$ the average number of packets correctly received in collisions of size n, we assume that the limit $C = \lim_{n\to\infty} C_n$ exists, as is usually the case with models of practical interest. It has been proved in [8] that the Aloha random access algorithm has a maximum stable throughput $\eta_0 = C$ in the multipacke channel.

Decentralized control strategies have been shown [11], [12], [19], [25], [30] to stabilize the slotted Aloha algorithm in the case of the usual collision channel, hence, it is reasonable to expect that when those strategies are used in the multipacket channel, the resulting throughput will be higher than η_0. We consider schemes of the form

$$p_n = F(S_n)$$

$$S_{n+1} = G(S_n, Z_n) \qquad (1)$$

where p_n is the retransmission probability in slot n, S_n is an estimate of the backlog X_n at the beginning of slot n, and Z_n is the feedback at the end of slot n. The number of new packets arriving during slot n, A_n is assumd to form a sequence of i.i.d. random variables with probability distribution $P[A_n = k] = \lambda_k (k \geq 0)$, such that the mean arrival rate $\lambda = \sum_{n=1}^{\infty} n\lambda_n$ is finite. Each of the A_{n-1} new packets that arrived during slot $n-1$ is transmitted in slot n with probability p_n.

As in the case of conventional channels, it is useful to study first the case of control with perfect state information where the value of the backlog is given to the users prior to the selection of the retransmission probability. To keep track of the exact value of the backlog, a central controller is usually necessary, which is an unreasonable requirement for most practical random access channels. However, the study of the perfect state information case allows us to determine an upper bound to the best throughput η_c achievable by any decentralized control of the form (1), and suggests a simple implementation. Those results are in turn helpful to derive control protocols in the case where the backlog is unknown. This is done in Section III where we consider a backlog estimate which is recursively updated using the binary feedback empty/nonempty. In addition, it is assumed throughout the paper that each station is informed when its packet is successfully received. It is proved that provided a certain condition on the reception matrix holds, the throughput achievable with this type of feedback is the same as the perfect state information throughput. This condition is verified for most multipoint-to-point channels of practical interest.

In a paper whose translation appeared only very recently [19] (after our work [7]), Mikhailov has derived sufficient conditions for stability and instability of two-dimensional Markov chains. Although this was meant to be used for decentralized control schemes in the usual collision channel, this approach is powerful enough to be applied to the multipacket channel. In Section IV we show by using Mikhailov's result that the scheme presented in Section III is stable under weaker assumptions. However, only a weaker form of stability can be proved in this way.

Manuscript received April 18, 1988; revised January 20, 1989 and May 7, 1989. Paper recommended by Associate Editor, X. R. Cao. This work was supported in part by the Office of Naval Research under Contracts N00014-87-K-0054, by the Army Research Office under Contract DAAL-03-87-K-0062, and by the New Jersey Commission on Science and Technology under Grant 85-990660-6.

The authors are with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544.

IEEE Log Number 8930788.

II. Control of the Multipacket Channel with Perfect State Information

In this section we assume that all the users know the value of X_n at the beginning of slot n, and we let the retransmission probability be a function of the exact value of the backlog, i.e., $p_n = F(X_n)$. In this ideal case, the system is much simpler to analyze than in the general case (1) since $(X_n)_{n\geq 0}$ is a homogeneous Markov chain. Our goal is to determine the optimal control function F^* that yields the largest ergodicity region, and the corresponding throughput, denoted by η_c. For instance, it is well known [4] that for the usual collision channel with the access rule in effect here, $F^*(X_n) = 1/X_n$ is the retransmission probability that minimizes the drift at each step, resulting in an ideal throughput of $\eta_c = e^{-1}$.

First note that all the results herein are valid provided that the backlog Markov chain $(X_n, S_n)_{n\geq 0}$ corresponding to a control (1) is irreducible and aperiodic. It can be easily checked that for both access rules considered in this paper (see below), as well as all the algorithms, a simple set of sufficient conditions for irreducibility and aperiodicity is

a) $\lambda_0 \neq 0$

b) $\lambda_0 + \sum_{n=1}^{\infty} \lambda_n \epsilon_{nn} < 1$

c) $\epsilon_{10} \neq 0$

which are analogous to the conditions for the open-loop system studied in [6]. The theorem below gives the best throughput possible for a control protocol (1).

Theorem 1: There exists a retransmission probability p_n^* that minimizes the expected backlog increase when the backlog is equal to n.

With such a retransmission probability, the system is stable for $\lambda < \eta_c$ and unstable for $\lambda > \eta_c$, with

$$\eta_c = \sup_{x\geq 0} e^{-x} \sum_{n=1}^{\infty} C_n \frac{x^n}{n!}.$$

Proof of Theorem 1: The proof is based on standard drift analysis techniques. $(X_n)_{t\geq 0}$ is a homogeneous Markov chain which evolves according to

$$X_{t+1} = X_t + A_t - \Sigma_t \tag{2}$$

where Σ_t is the number of packets successfully transmitted in slot t. The system is defined to be stable if $(X_t)_{t\geq 0}$ is ergodic and unstable otherwise. Let d_n be the drift of X_t at state n: $d_n = E[X_{t+1} - X_t | X_t = n]$. We have $0 \leq \Sigma_t \leq X_t$, and if we denote by p the retransmission probability used in slot t, then for $n \geq 1$, the probability of having k successes is given by

$$P[\Sigma_t = k | X_t = n] = \sum_{j=k}^{n} \binom{n}{j} p^j (1-p)^{n-j} \epsilon_{ij} \qquad (1 \leq k \leq n). \tag{3}$$

It then follows from (2) that the backlog drift at state $n \geq 1$ is given by

$$d_n = \lambda - \sum_{k=1}^{n} k \sum_{j=k}^{n} \binom{n}{j} p^j (1-p)^{n-j} \epsilon_{jk}$$

$$= \lambda - \sum_{j=1}^{n} \binom{n}{j} p^j (1-p)^{n-j} C_j \tag{4}$$

which becomes $d_n(p) = \lambda - t_n(p)$ if we define $t_n(p)$ to be the average number of successes given the backlog n and the retransmission probability p

$$t_n(p) = \sum_{j=1}^{n} \binom{n}{j} p^j (1-p)^{n-j} C_j. \tag{5}$$

Since $t_n(p)$ is a polynomial on the compact $[0,1]$, it achieves its maximum and we can define

$$p_n^* = \arg\max_{p\in[0,1]} t_n(p) = \arg\min_{p\in[0,1]} d_n(p).$$

We now proceed to compute the limit of the drift when the retransmission probability p_n^* is used. We show that

$$\lim_{n\to\infty} t_n(p_n^*) = \sup_{x\geq 0} e^{-x} \sum_{n=1}^{\infty} C_n \frac{x^n}{n!} = \sup_{x\geq 0} t(x). \tag{6}$$

Let us first assume that $C < +\infty$.

Property 1:

$$\lim_{x\to\infty} t(x) = C.$$

We have for $n > M$

$$|t(x) - C| \leq e^{-x} C + e^{-x} \sum_{n=1}^{M} \frac{x^n}{n!} |C_n - C| + \sum_{n=M+1}^{\infty} \frac{x^n}{n!} |C_n - C|. \tag{7}$$

Pick $\epsilon > 0$ and fix M such that $|C_n - C| < \epsilon$ for $n > M$. Then if B_c is an upper bound on the sequence $(C_n)_{n\geq 1}$, (7) yields

$$|t(x) - C| \leq e^{-x} C + 2B_c e^{-x} \sum_{n=1}^{M} \frac{x^n}{n!} + \epsilon$$

and the right-hand side of this last equation goes to zero as x goes to infinity.

Property 2: For all $\epsilon > 0$, there exists $A > 0$ such that for all $np > A$, $|t_n(p) - C| < \epsilon$. We have

$$|t_n(p) - C| \leq \sum_{j=1}^{n} \binom{n}{j} p^j (1-p)^{n-j} |C_j - C| + (1-p)^n C.$$

Choosing M as for Property 1 we get

$$|t_n(p) - C| \leq 2B_c \sum_{j=0}^{n} \binom{n}{j} p^j (1-p)^{n-j} + \epsilon.$$

Let us denote by R_n the random variable corresponding to the number of retransmissions in a slot given that the backlog is equal to n. We have

$$\sum_{j=0}^{M} \binom{n}{j} p^j (1-p)^{n-j} = P[R_n \leq M] \leq P\left[\left|\frac{R_n}{n} - p\right| > \frac{p}{2}\right]$$

or $np > 2M$. Then from the Chebyshev inequality

$$P[R_n \leq M] \leq \frac{4}{np} \tag{8}$$

and Property 2 follows.

Property 3: $t_n(x/n)$ converges uniformly to $t(x)$ on any compact $[0, A]$.

ix $\epsilon > 0$ and choose M such that $\sum_{j=M}^{\infty} A^j C_j/j! < \epsilon$. Then for $> M + 1$ and $x \in [0, A]$

$$t_n\left(\frac{x}{n}\right) - t(x) \Bigg| \le \sum_{j=1}^{M} A^j \frac{C_j}{j!} \left| e^{-x} \frac{n(n-1)\cdots(n-j+1)}{n^j} \left(1-\frac{x}{n}\right)^{n-j} \right| + 2\epsilon.$$

ince $\lim_{n\to\infty} n(n-1)\cdots(n-j+1)/n^j = 1$ for $1 \le j \le M$, t is enough to show that $(1 - x/n)^{n-j}$ converges uniformly to e^{-x} or $1 \le i \le M$. We have

$$1-\frac{x}{n}\Big)^{n-j} - e^{-x} \le e^{-x}[e^{xj/n} - 1] \le e^{AM/n} - 1. \tag{9}$$

)n the other hand, for $n > A$,

$$1-\frac{x}{n}\Big)^{n-j} - e^{-x} \ge \left(1-\frac{x}{n}\right)^{n} - e^{-x} \ge e^{-x}[e^{A+n\log(1-A/n)} - 1]$$

$$e^{n} \left(1-\frac{A}{n}\right)^{n} - 1 \tag{10}$$

ind uniform convergence follows from (9) and (10).

roperty 4: $t_n(x/n)$ converges uniformly to $t(x)$ for $x \ge 0$.

ix $\epsilon > 0$. From Properties 1 and 2 we can fix A such that:

) for all $np > A$, $|t_n(p) - C| < \epsilon$,

ii) for all $x > A$, $|t(x) - C| < \epsilon$.

hen we distinguish two cases. If $x \in [0, A]$, then from roperty 3 there exists N such that for all $n \ge N$, $|t_n(x/n) - t(x)| < \epsilon$. If on the other hand $x \in (A, +\infty)$, we have

$$t_n\left(\frac{x}{n}\right) - t(x) \Bigg| \le \left| t_n\left(\frac{x}{n}\right) - C \right| + |t(x) - C| \le 2\epsilon \tag{11}$$

rom i) and ii).

hus, we have shown that when C is finite, $t_n(x/n)$ converges iniformly to $t(x)$ for $x \ge 0$. It follows that $\lim_{n\to\infty} \sup_{x\ge 0} t_n(x/n) = \sup_{x\ge 0} t(x)$ and so (6) is proved.

inally, we show that (6) holds when $C = +\infty$. Choose Δ rbitrarily large and M such that $C_n > \Delta$ for $n > M$. Then for $n > M$

$$\left(1-\frac{x}{n}\right) \ge \Delta \sum_{j=M+1}^{n} \binom{n}{j} \left(\frac{x}{n}\right)^j \left(1-\frac{x}{n}\right)^{n-j} \ge \Delta(1 - P[R_n \le M]).$$

rom (8) $P[R_n \le M]$ is arbitrarily small for $nx/n = x$ large nough. Therefore, $\sup_{x\ge 0} t_n(x/n) = +\infty$ and $\lim_{n\to\infty} t_n(p_n^*) = -\infty$. Since it is clear that if $C = +\infty$, then $\sup_{x\ge 0} t(x) = +\infty$, 6) holds.

rom the equality $\lim_{n\to\infty} d_n(p_n^*) = \lambda - \sup_{x\ge 0} t(x)$ and Pakes emma in [21], it follows that if $\lim_{n\to\infty} C_n = +\infty$, then $\lim_{n\to\infty} t_n(p_n^*) = -\infty$, and the system is always stable, whereas if $\lim_{n\to\infty} C_n < +\infty$, then $(X_n)_{n\ge 0}$ is ergodic for $\lambda < \eta_c = \sup_{x\ge 0} t(x)$. lso, it is shown in the Appendix that Kaplan's condition holds or this system when the sequence $(C_n)_{n\ge 1}$ is bounded, thus from Kaplan's result [13], the backlog Markov chain is nonergodic when $\lambda > \eta_c$. □

t is intuitively obvious that no decentralized control algorithm of the form (1) can have a maximum stable throughput larger than η_c. The theorem below gives a rigorous proof of this fact and also nows that this throughput can be achieved with a control which is much simpler than p_n^*.

heorem 2: The best throughput achievable by a decentralized ontrol algorithm (1) is $\eta_c = \sup_{x\ge 0} e^{-x} \sum_{n=1}^{\infty} x^n/n! \, C_n$. If $\eta_c > C = \lim_{n\to\infty} C_n$, then there exists a constant $A > 0$ such that the ontrol $p_t = A/X_t$ for $X_t > A$ yields the optimal throughput η_c.

Proof of Theorem 2: To prove the first part of the theorem we use a result of [27] which is a generalization of Kaplan's heorem. If $p_t = F(S_t)$ and $S_{t+1} = G(S_t, Z_t)$, consider the Markov chain (X_t, S_t) and the Lyapunov function $V(n, s) = n$. Assume that $\lambda > \eta_c$. Then

$$E[V(X_{t+1}, S_{t+1}) - V(X_t, S_t) | X_t = n, S_t = s]$$

$$\lambda - \sum_{j=1}^{n} \binom{n}{j} F(s)^j (1-F(s))^{n-j} C_j$$

$$d_n(p_n^*) \ge \frac{\lambda - \eta_c}{2} \tag{12}$$

for all n large enough and all s. Therefore, the drift of V is strictly positive outside a finite subset of the state space. Since it is shown in the Appendix that the generalized Kaplan's condition is verified, it is enough to conclude that (X_t, S_t) is nonergodic. Hence, η_c is indeed the best throughput achievable by any decentralized control algorithm of the form (1).

To prove the second part of the theorem, we need the following property.

Property 5: If for all $x \ge 0$, $t(x) < \sup_{x\ge 0} t(x)$, then $\sup_{x\ge 0} t(x) = C$.

If $\sup_{x\ge 0} t(x) = +\infty$, it is easily seen that $C = +\infty$. If $\sup_{x\ge 0} t(x) < +\infty$, then $C < +\infty$. Consider a sequence $(x_n)_{n\ge 1}$ of nonnegative reals such that $\lim_{n\to\infty} t(x_n) = \sup_{x\ge 0} t(x)$. If $(x_n)_{n\ge 1}$ was bounded above by $K < +\infty$, we would have for all $n \ge 1$, $t(x_n) \le \sup_{x\in[0,K]} t(x)$, and in the limit $\sup_{x\ge 0} t(x) = \sup_{x\in[0,K]} t(x)$. Then there would exist $x_0 \in [0, K]$ such that $t(x_0) = \sup_{x\ge 0} t(x)$, which is a contradiction. Therefore, $(x_n)_{n\ge 1}$ is unbounded, and one can build a subsequence $(x_{n_k})_{k\ge 1}$ such that $\lim_{k\to\infty} x_{n_k} = +\infty$. We still have, of course, $\lim_{k\to\infty} t(x_{n_k}) = \sup_{x\ge 0} t(x)$, but on the other hand, we have $\lim_{k\to\infty} t(x_{n_k}) = \lim_{x\to\infty} t(x)$. From Property 1 in the proof of Theorem 1, $\lim_{x\to\infty} t(x) = C$ and Property 5 follows.

Thus, if $\eta_c > C$, then $t(x)$ achieves its supremum at some finite positive real A. Let us consider the control $p_t = A/X_t$ for $X_t \ge A$. (Note that the value of the retransmission probability is left unspecified for $X_t < A$ because it does not affect the throughput.) Then from (4) $d_n = \lambda - t_n(A/n)$, and from Property 3 in the proof of Theorem 1 $\lim_{n\to\infty} d_n = \lambda - t(A)$. Then it follows from [21] that $(X_t)_{t>0}$ is ergodic if $\lambda < t(A)$ and from [13] and the Appendix that $(X_t)_{t\ge 0}$ is nonergodic if $\lambda > t(A)$. Thus, the maximum stable throughput of the system is $t(A) = \sup_{x\ge 0} t(x) = \eta_c$. □

Note that the closed-loop throughput obtained in Theorems 1 and 2 can be interpreted as $\eta_c = \sup_{N\sim P(x), x>0} E[C_N]$, that is as the supremum over x of the expected value of C_N if N is a Poisson distributed random variable with mean x. Note that if we were to follow the popular approximation [1], [2], [10], [16], [18], [24], [26] that assumes that the number of transmissions in each slot, N, is Poisson distributed, and if we could choose any positive number as the mean of N by regulating the retransmission probability, the throughput would be equal to the average number of successes per slot, $E[C_N]$, maximized over the mean of N. As in the usual collision channel, a wrong analysis leads to a correct conclusion. Several examples are gathered in Table I (see [8] for details).

Probably the most important conclusion of this section is that in general it is not necessary to compute the exact value of p_n^*, which would require a large amount of on-line computations, and seriously hinder any application of Theorem 1 to the case where the backlog is unknown. Two cases may occur. If $t(x)$ does not attain its supremum, from Property 5 in the proof of Theorem 2, we have $\eta_c = \eta_0 = C$ (e.g., this happens in the model developed in [6] for mobile users with pairwise transmissions). In this case no throughput improvement can be achieved by varying the retransmission probability, and therefore it is enough to restrict attention to the open-loop strategy studied in [8]. On the other

TABLE I
OPEN-LOOP AND CLOSED-LOOP THROUGHPUTS FOR SEVERAL MULTIPACKET CHANNELS

	C_n	$\eta_0 = \lim_{n\to\infty} C_n$	$\eta_c = \sup_{A>0} e^{-A} \sum_{n=1}^{\infty} C_n \frac{A^n}{n!}$
conventional collision channel	$\begin{matrix}1 & n=1\\ 0 & n>1\end{matrix}$	0	e^{-1}
q-frequency frequency hopping [6]	$n(1-\frac{1}{q})^{n-1}$	0	$q\,e^{-1}$
mobile users with pairwise transmission [6]	1	1	1
capture - power discrimination [6]	$\begin{matrix}1 & n=1\\ \frac{1}{\beta^2} & n>1\end{matrix}$	$\frac{1}{\beta^2}$	$\frac{1}{\beta^2}+(1-\frac{1}{\beta^2})\exp(-\frac{\beta^2}{\beta^2-1})$
capture - timing discrimination [3]	$\begin{matrix}1 & n=1\\ (1-Q)^n & n>1\end{matrix}$	0	$\max_{A>0}\{(AQ-1)e^{-A}+e^{-AQ}\}$

and, if there exists A, $0 < A < +\infty$, such that $t(A) = \sup_{x\geq 0} t(x)$, then we have shown in the proof of Theorem 2 that the control $p_t = A/X_t$ for $X_t \geq A$ yields a maximum stable throughput $t(A) = \eta_c$, meaning that the system is optimal. Hence, only A has to be computed, and this can be done before starting the operation of the system.

Although in most practical applications $(C_n)_n \geq 1$ does have a limit, it is worth noticing that Theorem 1 can be generalized to the case where C does not exist. It can be shown [9] that if the drift is minimized at each step, then the system is stable for $\lambda < \sup_{x\geq 0} t(x)$ and unstable for $\lambda > \sup_{x\geq 0} t(x) + \lim_{n\to\infty} \sup C_n - \lim_{n\to\infty} \inf C_n$. As in the open-loop system when $(C_n)_{n\geq 1}$ does not have a limit, nothing more can be said about the throughput without further information on the sequence $(C_n)_{n\geq 1}$. But the main drawback in such a case is that there may not exist any control $p_n = A/X_n$ that yields the optimal throughput.

The access rule for new packets that we have been considering so far is usually referred to as delayed first transmission (DFT). With this access rule, newly arrived packets are treated exactly in the same way as backlogged packets. Let us now examine what happens when on the contrary an immediate first transmission (IFT) rule is used, that is when new packets are transmitted with probability one in the slot immediately following their arrival. It has been proved in [8] that the open-loop throughput is the same for both first transmission rules. The closed-loop throughput on the other hand depends on the access rule. For instance, it is well known [4] that for the usual collision channel in the IFT case, the optimal retransmission probability is $p_n^* = \lambda_0 - \lambda_1/\lambda_0 n - \lambda_1$, yielding an optimal throughput $\lambda_0 e^{\lambda_1/\lambda_0} e^{-1}$, in contrast to the throughput $\eta_c = e^{-1}$ for the DFT case. In the multipacket channel with the IFT rule, the optimal throughput depends not only on the mean but on the whole distribution of new packet arrivals. Interestingly enough, it can be proved that both throughputs coincide when the new packet arrivals are Poisson distributed. Still with the same method as in the proof of Theorem 1, it can be easily shown that there exists a retransmission probability that minimizes the drift d_n at state n. With such a retransmission probability, the system with IFT rule is stable for $\lambda < \sup_{x\geq 0} T(x)$ and unstable for $\lambda > \sup_{x\geq 0} T(x)$, with $T(x) = e^{-x} \sum_{n=0}^{\infty} x^n/n! \sum_{j=0}^{\infty} \lambda_j C_{n+j}$, where we have defined $C_0 = 0$ for notational convenience. It can also be proved that a control of the form $p_n = A/X_n$ yields a maximum stable throughput $T(A)$. Since $\sup_{x\geq 0} T(x)$ depends on the whole new packet arrival distribution $(\lambda_n)_{n\geq 0}$, this result is not as conclusive as in the DFT case. This is because the stability region $\lambda < \sup_{x\geq 0} T(x)$ is actually given in the form of an implicit equation in λ, which cannot be solved in general without further specifications on the distribution $(\lambda_n)_{n\geq 0}$. For instance, this stability region could be empty. Consider, for example, the usual collision channel with possibly some added noise $0 < C_1 \leq 1$ and $C_n = 0$ for $n \geq 2$. Then $T(x) = C_1 e^{-x}(\lambda_1 + \lambda_0 x)$ and $T'(x) = C_1 e^{-x}(\lambda_0 - \lambda_1 - \lambda_0 x)$. Therefore, for any distribution such that $\lambda_0 < \lambda_1$, $T(x)$ is maximum at $T(0) + C_1\lambda_1$, and the stability region is empty since $C_1\lambda_1 \leq \lambda_1 \leq \lambda$. Note that in this sense, the immediate first transmission does not perform as well as the delayed first transmission with which the system can always be stabilized.

If there are solutions to $\lambda < \sup_{x\geq 0} T(x)$, then the best throughput achievable by the class of algorithms in (1) is $\nu_c = \sup\{\lambda : \lambda < \sup_{x\geq 0} T(x)\}$. This is what happens, for instance, when the new packet arrivals are Poisson distributed.

Theorem 3: If the new packet arrivals are Poisson distributed, the best throughput achievable with an IFT rule is the same as in the DFT case, $\nu_c = \sup_{x\geq 0} t(x)$.

Proof of Theorem 3: If $\lim_{n\to\infty} C_n = +\infty$, then $\eta_c = \nu_c = +\infty$. Assume now that $C < +\infty$. We get

$$T(x) = e^{-(x+\lambda)} \sum_{n=0}^{\infty} \frac{x^n}{n!} \sum_{k=0}^{\infty} \sum_{k!}^{\lambda^k} C_{n+k}$$

$$= e^{-(x+\lambda)} \sum_{n=1}^{\infty} \frac{C_n}{n!} (x+\lambda)^n. \tag{13}$$

Thus, in this case, $T(x)$ depends only on λ, and to clarify the proof below, we denote it by $T_\lambda(x)$

$$T_\lambda(x) = t(x+\lambda). \tag{14}$$

Assume that $t(x)$ does not achieve its supremum. Then from Property 5 in the proof of Theorem 2, we have $\eta_c = C = \lim_{x\to\infty} t(x)$. It follows from (14) that for any $\lambda > 0$, $\lim_{x\to\infty} T_\lambda(x) = C$. Therefore, for all $\lambda > 0$, $\sup_{x\geq 0} T_\lambda(x) \geq C$. Hence, for all $\lambda > 0$, $\sup_{x\geq 0} T_\lambda(x) = \sup_{x\geq 0} t(x)$, and by definition of ν_c, we finally get $\nu_c = \sup_{x\geq 0} t(x)$. Note that T_λ does not achieve its supremum, in the sense that if there existed $\lambda \in (0, \nu_c)$ and $x_\lambda \geq 0$ such that $\nu_c = T_\lambda(x_\lambda)$, we would have $\sup_{x\geq 0} t(x) = t(\lambda + x_\lambda)$.

Assume now that $t(x)$ does achieve its supremum, there exists $x_0 \geq 0$ such that $\sup_{x\geq 0} t(x) = t(x_0)$. Then for all λ in $[0, x_0]$: $T_\lambda(x_0 - \lambda) = \sup_{x\geq 0} t(x) \geq \sup_{x\geq 0} T_\lambda(x)$. Thus, for all $\lambda \in [0, x_0]$

$$\sup_{x\geq 0} T_\lambda(x) = \sup_{x\geq 0} t(x) = T_\lambda(x_0 - \lambda). \tag{15}$$

We have for all $x \geq 0$ $t(x) \leq x$, therefore $\sup_{x\geq 0} t(x) \leq x_0$. Together with (15), it follows that for all $\lambda \in (0, \sup_{x\geq 0} t(x))$, $\lambda < \sup_{x\geq 0} T_\lambda(x)$, and therefore $\nu_c \geq \sup_{x\geq 0} t(x) = \eta_c$. Since from (14) $\sup_{x\geq 0} T_\lambda(x) \leq \sup_{x\geq 0} t(x) = \eta_c$ for all λ, we get $\nu_c \leq \eta_c$ and finally $\nu_c = \eta_c = \sup_{x\geq 0} t(x)$. Note that from (14), T_λ reaches its supremum too, since for all $\lambda < \nu_c$, there exists $x_\lambda \geq 0$ such that $T_\lambda(x_\lambda) = \nu_c$.

Note that we have also shown in this proof that $T(x)$ reaches its supremum iff $t(x)$ does, which means that η_c can be achieved with a control of the form $p_n = A/X_n$ iff ν_c can. □

III. Optimal Control for the Multipacket Channel

It is assumed from now on that the users do not have access to the value of the backlog, so the problem becomes one of control of the Markov chain with partial state information provided by the channel feedback. We build a backlog estimate S_t with feedback which is such that $Z_t = 0$ if slot t was empty, and $Z_t = 0$ otherwise. The results of the previous section strongly suggest that we should use as a retransmission probability $p_t = A/S_t$, where A is a point at which $t(x)$ achieves its supremum (according to Property 5, A is assumed to be finite). We show that the resulting control algorithm achieves the optimal maximum stable throughput η_c. This holds provided that the following assumption on the reception matrix is verified.

C0: There exists $\theta > 0$ and B such that for all $n \geq 1$, $\Sigma_{k=1}^{n} e^{\theta k} \nu_k \leq B$.

The purpose of condition C0 is to bound the probability of having large numbers of simultaneous successes. Unbounded numbers of successes per slot are difficult to deal with because they may result in very large instantaneous errors in the backlog estimate. Note that condition C0 is likely to hold in most multipoint-to-point channels because of practical limitations on the receiver capabilities, and that it is verified for all the examples in Table I.

Theorem 4: Assume that there exists $A \in (0, +\infty)$ such that $\nu(A) = \sup_{x \geq 0} \nu(x)$, that the new packet arrivals $(A_t)_{t \geq 0}$ are exponential type[1], and that condition C0 holds. If $\alpha < 0$ and $\beta < 1$ verify the following two conditions[2]:

C1: $\beta > \lambda$

C2: $\beta(1 - e^{-A}) + \eta_c - \lambda + \alpha e^{-A} = 0$

then the control algorithm (cf. the control laws proposed in [15], [19], and [25])

$$\nu_t = \frac{A}{S_t}$$

$$S_{t+1} = \max \{A, S_t + \alpha I(Z_t = 0) + \beta I(Z_t = 0)\}$$

has maximum stable throughput equal to η_c.

Proof of Theorem 4: The proof is based on the method developed in [30]. The idea is to use the properties of the homogeneous two-dimensional vector Markov chain of the backlog and its estimate $M_t = (X_t, S_t)$ to build a Lyapunov function whose drift is negative in the first quadrant of the (n, s) plane when $\lambda < \eta_c$. It turns out that this fails to hold in two cones of the state space, but it can be proved that the J-step drift of the Lyapunov function is negative for some integer J, and that this is enough to ensure that M_t is geometrically ergodic. It follows from Theorem 2 that M_t is nonergodic if $\lambda > \eta_c$. For substantial portions of the proof, the reader is referred to [9] because of space limitations.

Denote by $\tilde{X}_t = S_t - X_t$ the error in the backlog estimate. The first part of the proof mainly consists of computing and approximating the drifts of X_t and $\tilde{X}_t$ which are the basic building blocks for the Lyapunov function.

Denote by $c(n, s) = E[X_{t+1} - X_t | M_t = (n, s)]$ the backlog drift at state (n, s), and by $d(n, s) = E[\tilde{X}_{t+1} - \tilde{X}_t | M_t = (n, s)]$ the drift of the backlog error. For technical reasons, what we most often use in the proof are the truncated drifts, which correspond to the value of the drifts restricted to those paths where the variation in the backlog is bounded by some integer J, that is $c(n, s, J) = E[(X_{t+1} - X_t) I(|X_{t+1} - X_t| \leq J) | M_t = (n, s)]$ and $d(n, s, J) = E[(X_{t+1} - \tilde{X}_t) I(|X_{t+1} - X_t| \leq J) | M_t = (n, s)]$. Clearly, these truncated drifts will be good approximations of $c(n, s)$ and $d(n, s)$, respectively, when J is large. It will turn out that the drifts depend primarily on the ratio $x = n/s$ for large values of n or s. Thus, it is convenient to define the following two regions in the (n, s) plane:

$$\mathcal{C}(\lambda_0, \lambda_1) = \{(n, s) : n \geq 0, s \geq 0, 1 + \lambda_0 \leq \frac{n}{s} \leq 1 + \lambda_1\}$$

$$\mathcal{I}_M = \{(n, s) : n \geq M \text{ or } s \geq M\}$$

where λ_0 and λ_1 are such that $-\infty \leq \lambda_0 \leq \lambda_1 \leq +\infty$. The aim of the first part of the proof is to show Proposition 1 below which summarizes all the properties of the drifts that are needed for our purposes (see Fig. 1).

[1] A_t is exponential type if there exists $d > 0$ such that $E[e^{dA_t}]$ is finite. For instance, this is true if A_t is Poisson distributed.

[2] Conditions C1 and C2 define half a straight line in the plane, and therefore an infinite number of possible estimation schemes, all of them yielding the same throughput.

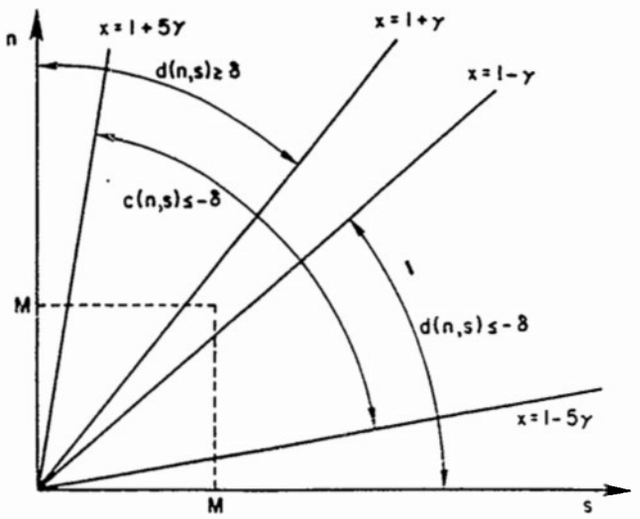

Fig. 1. Drift properties (Proposition 1).

Proposition 1: There exist $\gamma \in (0, 1/5)$, $\delta > 0$, and an integer $J_0 > 0$ such that for all $J \geq J_0$:

i) for all $(n, s) \in C(-5\gamma, 5\gamma) \cap U_M$, $c(n, s) \leq -\delta + \nu(J)$ and $c(n, s, J) \leq -\delta + \nu(J)$;

ii) for all $(n, s) \in C(-\infty, -\gamma) \cap U_M$, $d(n, s) \leq -\delta$ and $d(n, s, J) \leq -\delta + \nu(J)$;

iii) for all $(n, s) \in C(\gamma, +\infty) \cap U_M$, $d(n, s) \geq \delta$ and $d(n, s, J) \geq \delta - \nu(J)$

where $\nu(J)$ is a nonnegative function which goes to zero a J goes to infinity.

The detailed proof of Proposition 1 can be found in [9]. After computing the value of the drifts

$$c(0, s) = \lambda \tag{16a}$$

$$c(n, s) = \lambda - \sum_{j=1}^{n} \binom{n}{j} \left(\frac{A}{s}\right)^j \left(1 - \frac{A}{s}\right)^{n-j} C_j \qquad (n \geq 1) \tag{16b}$$

$$d(0, s) = \max \{A - s, \alpha\} - \lambda \tag{17a}$$

$$d(n, s) = \beta - \lambda + (\max \{A - s, \alpha\} - \beta) \left(1 - \frac{A}{s}\right)^n + \sum_{j=1}^{n} \binom{n}{j} \left(\frac{A}{s}\right)^j \left(1 - \frac{A}{s}\right)^{n-j} C_j \qquad (n \geq 1) \tag{17b}$$

we work out upper and lower bounds by truncating the sums (16) and (17) to a fixed number of terms, and then we approximate those bounds as a function of the sole variable n/s. The main idea is that the dynamic behavior of the Markov vector $M_t = (X_t, S_t)$ depends essentially on the ratio X_t/S_t. For instance, if x is nearly equal to 1, the backlog estimate is close to its ideal value, and we should have $c(n, s) < 0$ since the backlog drift is negative in the perfect state information case. Also, a well-behaved estimate should be such that if $x < 1$, then the error $s - n$ is positive, and therefore should have a negative drift $d(n, s) < 0$ (see [15]). In the same way, we expect to have $d(n, s) > 0$ for $x > 1$.

Let us define the following Lyapunov function:

$$V(n, s) = \max \left\{n, \frac{1 + 3\gamma}{3\gamma}(n - s), \frac{1 - 3\gamma}{3\gamma}(s - n)\right\}$$

where the constants have been chosen so that V is continuous. $V(n, s)$ is equal to the first, second, and third term inside the bracket when (n, s) is in $C(-3\gamma, 3\gamma)$, $C(3\gamma, +\infty)$, and $C(-\infty, -3\gamma)$, respectively. Notice that V is defined so as to take the best advantage of the drift properties listed in Proposition 1. For

instance, when $V(n, s)$ is equal to n, then the Markov chain M_t belongs to $C(-3\gamma, 3\gamma)$ which is included in $C(-5\gamma, 5\gamma)$ where the backlog drift is negative provided that either n or s is sufficiently large. Similar comments can be made about the other two regions. Unfortunately, this does not enable us to conclude that the drift of the Lyapunov function is negative in U_M because M_{t+1} may well be in a different region than M_t. However, this change of region becomes unlikely if we exclude a small zone around the lines $x = 1 \pm 3\gamma$ where V changes definition and indeed the second part of this proof consists of showing that the Lyapunov function has a negative drift in the remainder of the state space.

Proposition 2: There exist $M_0 \geq 0$ and $\delta_0 > 0$ such that for all $N \geq M_0$ and for all $(n, s) \in U_N \cap [C(-\infty, -4\gamma) \cup C(-2\gamma, 2\gamma) \cup C(4\gamma, \infty)]$,

$$E[V(M_{t+1}) - V(M_t) | M_t = (n, s)] < -\delta_0.$$

Proof of Proposition 2: We consider separately likely and unlikely events

$$\begin{aligned} &E[V(M_{t+1}) - V(M_t) | M_t = (n, s)] \\ &\quad = E[(V(M_{t+1}) - V(M_t)) I(|A_t - \Sigma_t| \leq J) | M_t = (n, s)] \\ &\qquad + E[(V(M_{t+1}) - V(M_t)) I(|A_t - \Sigma_t| > J) | M_t = (n, s)]. \end{aligned} \quad (18)$$

We start by showing that the first term, which corresponds to likely events, is negative when J is large by using the properties of the truncated drifts from Proposition 1 and a simple geometric result. The lemma below, whose proof is in [9], gives a measure of how much a cone $C(\lambda_0, \lambda_1)$ expands if each of its points is allowed to move of some distance that cannot exceed B in absolute value along each axis.

Lemma: Consider $\gamma > 0$, $B > 0$, and $\gamma - 1 < \lambda_0 < \lambda_1 < +\infty$; and assume that $|n - n'| \leq B$, $|s - s'| \leq B$, and $Q \geq B/\gamma (1 + |\lambda_1|)(\lambda_1 + 2 + \gamma)$. Then:

1) $(n, s) \in C(\lambda_0, \infty) \cap U_Q$
$\Rightarrow (n', s') \in C(\lambda_0 - \gamma, \infty) \cap U_{Q-B}$

2) $(n, s) \in C(-\infty, \lambda_1) \cap U_Q$
$\Rightarrow (n', s') \in C(-\infty, \lambda_1 + \gamma) \cap U_{Q-B}$

3) $(n, s) \in C(\lambda_0, \lambda_1) \cap U_Q$
$\Rightarrow (n', s') \in C(\lambda_0 - \gamma, \lambda_1 + \gamma) \cap U_{Q-B}$.

Set $B(J) = \max\{J, |\alpha| + \beta\}$, and define $Q(J)$ to be any real such that $Q(J) \geq \max\{B(J) + M, B(J)/\gamma (1 + 4\gamma)(2 + 3\gamma)\}$. We have $|S_{t+1} - S_t| \leq |\alpha| + \beta \leq B(J)$, and if $|A_t - \Sigma_t| \leq J$, then $|X_{t+1} - X_t| \leq J \leq B(J)$. From the lemma, $Q(J)$ is such that

$$M_t \in C(-2\gamma, 2\gamma) \cap U_{Q(J)} \Rightarrow M_{t+1} \in C(-3\gamma, 3\gamma) \cap U_M \quad (19)$$

$$M_t \in C(4\gamma, \infty) \cap U_{Q(J)} \Rightarrow M_{t+1} \in C(3\gamma, \infty) \cap U_M \quad (20)$$

$$M_t \in C(-\infty, -4\gamma) \cap U_{Q(J)} \Rightarrow M_{t+1} \in C(-\infty, -3\gamma) \cap U_M \quad (21)$$

where M has been defined in Proposition 1. Assume, for instance that M_t belongs to $C(-2\gamma, 2\gamma) \cap U_{Q(J)}$. From (19), $M_{t+1} \in C(-3\gamma, 3\gamma) \cap U_M \cap C(-5\gamma, 5\gamma) \cap U_M$. Hence, if $J \geq J_0$, we can apply Proposition 1 i):

$$E[(V(M_{t+1}) - V(M_t)) I(|A_t - \Sigma_t| \leq J) | M_t = (n, s)] \leq -\delta + \nu(J).$$

If M_t belongs to the other two regions, $C(4\gamma, \infty) \cap U_{Q(J)}$ or $C(-\infty, -4\gamma) \cap U_{Q(J)}$, a similar argument holds, using Proposition 1 iii) and ii), respectively, along with (20) and (21). It follows that for all $J \geq J_0$ and for all $(n, s) \in U_{Q(J)} \cap [C(-\infty, -4\gamma) \cup C(-2\gamma, 2\gamma) \cup C(4\gamma, \infty)]$

$$E[(V(M_{t+1}) - V(M_t)) I(|A_t - \Sigma_t| \leq J) | M_t = (n, s)] \leq -\delta_1 + \nu_1(J) \quad (22)$$

with $\delta_1 = \min\{1, 1 - 3\gamma/3\gamma\}\delta$ and $\nu_1(J) = \nu(J)\, 1 + 3\gamma/3\gamma$.

To deal with the second term on the right-hand side of (18), we consider the further decomposition

$$\begin{aligned} &E[(V(M_{t+1}) - V(M_t)) I(|A_t - \Sigma_t| > J) | M_t = (n, s)] \\ &\quad = E[(V(M_{t+1}) - V(M_t)) I(|A_t > \Sigma_t + J) | M_t = (n, s)] \\ &\qquad + E[(V(M_{t+1}) - V(M_t)) I(\Sigma_t > A_t + J) | M_t = (n, s)]. \end{aligned} \quad (23)$$

Let us denote by $T_1(n, s, J)$ and $T_2(n, s, J)$ the two terms on the right-hand side of (23). The first term $T_1(n, s, J)$ corresponds to a case where the variation in the backlog is bounded below, and can be shown to vanish as J increases by using the sole fact that the mean arrival rate λ is finite. Consider now $T_2(n, s, J)$. If $M_t = (n, s)$ belongs to a region such that $x = n/s > x_0$, then x_0 can be chosen large enough so that if M_{t+1} belongs to $C(-\infty, -3\gamma)$, then the error in the backlog estimate which results from the large number of successes just compensates the initial error $n - s \gg 0$. On the other hand, when M_t belongs to any region such that x is bounded above, then $E[\Sigma_t I(\Sigma_t > J) | M_t = (n, s)]$ goes to zero uniformly in (n, s) and $T_2(n, s, J)$ can be dealt with by using the following rather crude bound for the variation of V:

$$\begin{aligned} |V(M_{t+1}) - V(M_t)| &\leq \max\left\{1, \frac{1 + 3\gamma}{3\gamma}, \frac{1 - 3\gamma}{3\gamma}\right\} \\ &\quad \cdot (|\alpha| + \beta + |A_t - \Sigma_t|) \leq R(1 + |A_t - \Sigma_t|) \end{aligned} \quad (24)$$

where R is some positive constant. It is shown in [9] that

$$E[(V(M_{t+1}) - V(M_t)) I(|A_t - \Sigma_t| > J) | M_t = (n, s)] \leq \nu_2(J) + \epsilon_J(n, s) \quad (25)$$

where $\lim_{J \to \infty} \nu_2(J) = 0$, and $\epsilon_J(n, s)$ is a nonnegative function that depends on J, and goes to zero as either n or s goes to infinity.

By using (22), (25), and the decomposition (18), we get the desired result that the drift of V is negative in this part of the state space: fix an integer $J_{\min}$ such that $J_{\min} \geq J_0$ and that for all $J \geq J_{\min}$, $\nu_1(J) + \nu_2(J) \leq \delta_1/3$. Then from (22) and (25), we have for all $(n, s) \in U_{Q(J_{\min})} \cap [C(-\infty, -4\gamma) \cup C(-2\gamma, 2\gamma) \cup C(4\gamma, \infty)]$,

$$E[V(M_{t+1}) - V(M_t) | M_t = (n, s)] \leq -\frac{2}{3}\delta_1 + \epsilon_{J_{\min}}(n, s).$$

Then we can choose an $M_0 > Q(J_{\min})$ which is large enough so that $\epsilon_{J_{\min}}(n, s) < \delta_1/3$ for all (n, s) in U_{M_0}. □

This concludes the second part of the proof. Unfortunately, it is not always true that the drift of V is negative outside a finite subset of the state space. For instance, we have proved that in the case of the usual collision channel with Poisson new packet arrivals, there exist constants $B_{ex} > 0$ and M_{ex} such that for all $(n, s) \in U_{M_{ex}}$ for which $x = 1 \pm 3\gamma$, and for all α and β verifying C1 and C2, $E[V(M_{t+1}) - V(M_t) | M_t = (n, s)] > B_{ex}$. However, discontinuities around the lines $x = 1 \pm 3\gamma$ cancel out when one waits long enough, and in the last part of this proof we show that the J-step drift of V, $E[V(M_{t+J}) - V(M_t) | M_t = (n, s)]$ is negative for some integer J.

Proposition 3: There exist $J_f > 0$, $\rho > 0$, and $M_f > 0$ such

that for all $(n, s) \in U_{M_f}$

$$E[V(M_{t+J}) - V(M_t) | M_t = (n, s)] \leq -\rho.$$

Proof of Proposition: One of the main problems in dealing with the J-step drift of V is to control the changes of regions between M_t and M_{t+J}. To this end, we define the stopping time

$$\tau_J = \min \left\{ s \geq 0, \left| \sum_{k=0}^{s} (A_{t+k} - \Sigma_{t+k}) \right| > J^3 \right\}.$$

If $\tau_J \geq J$, then for $1 \leq k \leq J$, $|X_{t+k} - X_t| \leq J^3$ and $|S_{t+k} - S_t| \leq J(|\alpha| + \beta)$. Thus, if we define $B'(J) = \max \{J(|\alpha| + \beta), J^3\}$, and $Q'(J)$ to be any integer such that $Q'(J) \geq B'(J) + \max \{M_0, M\}$ and $Q'(J) \geq 2B'(J)/\gamma(1 + 9/2\gamma)(5\gamma + 2)$, then, still assuming that $\tau_J \geq J$, we get from the lemma for $0 \leq k \leq J$

$$M_t \in C\left(-\infty, -4\gamma - \frac{\gamma}{2}\right) \cap U_{Q'(J)} \Rightarrow M_{t+k} \in C(-\infty, -4\gamma) \cap U_{M_0} \quad (26)$$

$$M_t \in C\left(-2\gamma + \frac{\gamma}{2}, 2\gamma - \frac{\gamma}{2}\right) \cap U_{Q'(J)} \Rightarrow M_{t+k} \in C(-2\gamma, 2\gamma) \cap U_{M_0} \quad (27)$$

$$M_t \in C\left(4\gamma + \frac{\gamma}{2}, \infty\right) \cap U_{Q'(J)} \Rightarrow M_{t+k} \in C(4\gamma, \infty) \cap U_{M_0} \quad (28)$$

$$M_t \in C\left(-4\gamma - \frac{\gamma}{2}, -2\gamma + \frac{\gamma}{2}\right) \cap U_{Q'(J)} \Rightarrow M_{t+k} \in C(-5\gamma, -\gamma) \cap U_M \quad (29)$$

$$M_t \in C\left(2\gamma - \frac{\gamma}{2}, 4\gamma + \frac{\gamma}{2}\right) \cap U_{Q'(J)} \Rightarrow M_{t+k} \in C(\gamma, 5\gamma) \cap U_M. \quad (30)$$

In other words, we have partitioned the plane into two zones

$$Z_N = C\left(-\infty, -4\gamma - \frac{\gamma}{2}\right) \cup C\left(-2\gamma + \frac{\gamma}{2}, 2\gamma - \frac{\gamma}{2}\right) \cup C\left(4\gamma + \frac{\gamma}{2}, \infty\right),$$

and

$$Z_P = C\left(-4\gamma - \frac{\gamma}{2}, -2\gamma + \frac{\gamma}{2}\right) \cup C\left(2\gamma - \frac{\gamma}{2}, 4\gamma + \frac{\gamma}{2}\right).$$

Then we have chosen $Q'(J)$ such that if M_t belongs to Z_N which is slightly smaller than the region in which the drift of the Lyapunov function is negative, and if $\tau_J \geq J$, then the Markov chain remains in the region in which Proposition 2 applies up to time $t + J$ (see (26)–(28) and Fig. 2). $Q'(J)$ is also such that if M_t is in Z_P and if $\tau_J \geq J$, then up to time $t + J$ the chain stays in a region such that two out of the three properties of Proposition 1 hold at each step (see (29), (30), and Fig. 3).

We start by showing that the J-step drift of V is negative at (n, s) when (n, s) belongs to Z_N. We decompose the J-step drift of V as follows:

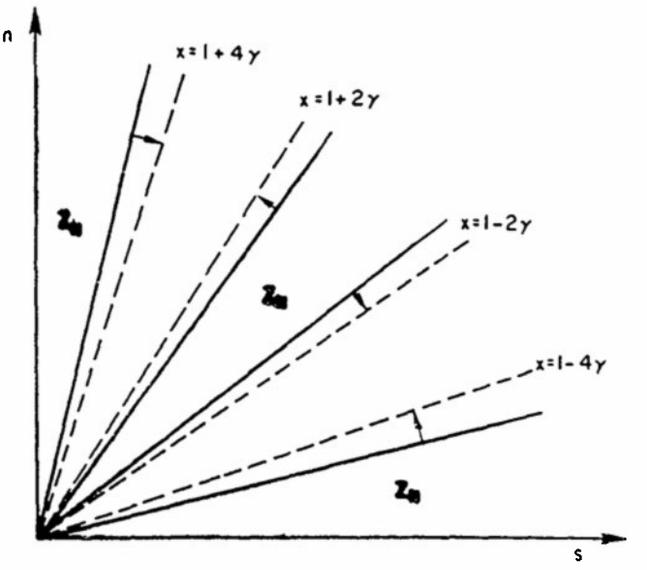

Fig. 2. If $M_t \in Z_N \cap U_{Q'(J)}$ and if $\tau_J \geq J$, then M_{t+1} belongs to the region where the drift of V is negative.

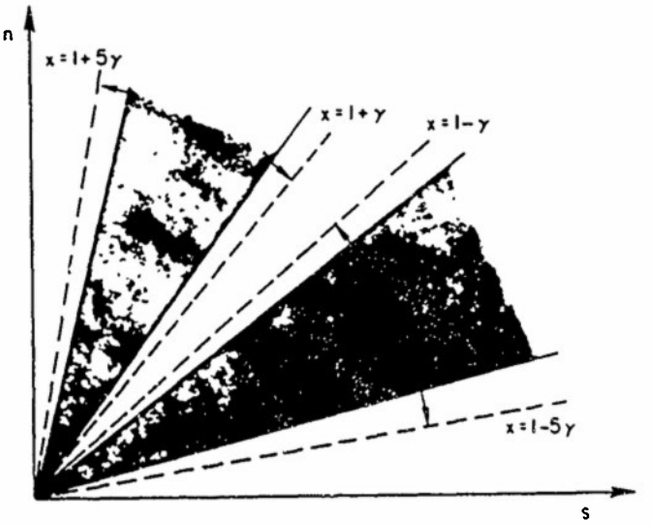

Fig. 3. If $M_t \in Z_P \cap U_{Q'(J)}$ and if $\tau_J \geq J$, then M_{t+1} belongs to a region where two properties of Proposition 1 hold.

$$E[V(M_{t+J}) - V(M_t) | M_t = (n, s)] = \sum_{k=0}^{J-1} E[E[V(M_{t+k+1}) - V(M_{t+k})] | M_{t+k}] \cdot I(\tau_J \geq J) | M_t = (n, s)] + \sum_{k=0}^{J-1} E[E[V(M_{t+k+1}) - V(M_{t+k}) | M_{t+k}] I(\tau_J < J) | M_t = (n, s)]. \quad (31)$$

Denote by $U_1(J, n, s)$ and $U_2(J, n, s)$ the two sums on the right-hand side of (31). If $\tau_J \geq J$, then (26)–(28) hold, and therefore we can apply Proposition 2

$$U_1(J, n, s) \leq -J\delta_0 P[\tau_J \geq J | M_t = (n, s)]. \quad (32)$$

Let us now show that $\tau_J < J$ is indeed an unlikely event, the

probability of which goes to zero as $1/J$ uniformly in (n, s)

$$P[\tau_J < J | M_t = (n, s)]$$

$$\leq \sum_{l=0}^{J-1} P\left[\left|\sum_{l=0}^{k} (A_{t+l} - \Sigma_{t+l})\right| > J^3 | M_t = (n, s)\right]$$

$$\leq \sum_{l=0}^{J-1} P\left[\sum_{l=0}^{k} A_{t+l} > J^3\right] + \sum_{l=0}^{J-1} P\left[\sum_{l=0}^{k} \Sigma_{t+l} > J^3 | M_t = (n, s)\right].$$

From Markov's inequality we have

$$P[\tau_J < J | M_t = (n, s)] \leq \frac{1}{J^3} \sum_{l=0}^{J-1} (k+1)\lambda + \frac{1}{J^3} \sum_{l=0}^{J-1} \sum_{l=0}^{k} E[\Sigma_{t+l} | M_t = (n, s)].$$

Denoting by B_η an upper bound on the sequence $t_n(p_n^*)$, it follows from Section II that $E[\Sigma_{t+l} | M_t = (n, s)] = E[E[\Sigma_{t+l} | M_{t+1}] | M_t = (n, s)] \leq B_\eta$, so we get

$$P[\tau_J < J | M_t = (n, s)] \leq \frac{\lambda + B_\eta}{J^2} \frac{J+1}{J^2} \leq \frac{B_\tau}{J} \tag{33}$$

where B_τ is some positive constant. From (24), it is easy to check that the drift of V is bounded by some positive constant B_V, so that

$$J_J(J, n, s) \leq J B_V P[\tau_J < J | M_t = (n, s)]. \tag{34}$$

Considering (31), (32), (33), and (34), we get

$$E[V(M_{t+J}) - V(M_t) | M_t = (n, s)] \leq -\delta_0 J + (B_V + \delta_0) B_\tau.$$

Therefore, there exist constants $\mu_1 > 0$ and $J_1 > 0$ such that for all $J \geq J_1$ and for all $(n, s) \in U_{Q^c(J)} \cap Z_N$,

$$E[V(M_{t+J}) - V(M_t) | M_t = (n, s)] \leq -J\mu_1. \tag{35}$$

We now proceed to show that the J-step drift of the Lyapunov function is negative in the remaining part of the state space Z_P consisting of the two cones around $x = 1 \pm 3\gamma$. This is done in two steps. We first show that the J-step drift of V restricted to likely events $\{\tau_J \geq J\}$ goes to $-\infty$ as J increases, and then we prove that the J-step drift of V restricted to unlikely events $\{\tau_J < J\}$ is bounded above independent of J.

Assume, for instance, that $(n, s) \in C(\gamma - \gamma/2, 4\gamma + \gamma/2) \cap U_{Q^c(J)}$. The difficulty here is that V can take two possible values, and therefore Proposition 1 cannot be used directly. If $\tau_J \geq J$, then from (30) $M_{t+k} \in C(\gamma, 5\gamma) \cap U_M$ for $0 \leq k \leq J$, so that $V(M_{t+k}) = \max\{X_{t+k}, (1 + 3\gamma)/3\gamma(X_{t+k} - S_{t+k})\}$. Therefore,

$$E[(V(M_{t+J}) - V(M_t)) I(\tau_J \geq J) | M_t = (n, s)]$$

$$\leq E\left[\max\left\{X_{t+J}, \frac{1+3\gamma}{3\gamma}(X_{t+J} - S_{t+J})\right\} \cdot I(\tau_J \geq J) | M_t = (n, s)\right]$$

$$- E\left[\max\left\{X_t, \frac{1+3\gamma}{3\gamma}(X_t - S_t)\right\} \cdot I(\tau_J \geq J) | M_t = (n, s)\right]$$

$$\leq E\left[\max\left\{X_{t+J} - X_t, \frac{1+3\gamma}{3\gamma}(-\tilde{X}_{t+J} + \tilde{X}_t)\right\} \cdot I(\tau_J \geq J) | M_t = (n, s)\right]$$

since $\max\{a, b\} - \max\{c, d\} \leq \max\{a - c, b - d\}$. Then using the fact that $\max\{a, b\} \leq \max\{0, a + f\} + \max\{0, b + f\} - f$ for $f \geq 0$, we get

$$E[(V(M_{t+J}) - V(M_t)) I(\tau_J \geq J) | M_t = (n, s)]$$

$$\leq E\left[\max\left\{0, X_{t+J} - X_t + \delta_1 \frac{J}{2}\right\} \cdot I(\tau_J \geq J) | M_t = (n, s)\right]$$

$$+ E\left[\max\left\{0, \frac{1+3\gamma}{3\gamma}(-\tilde{X}_{t+J} + \tilde{X}_t) + \delta_1 \frac{J}{2}\right\} \cdot I(\tau_J \geq J) | M_t = (n, s)\right]$$

$$- E\left[\delta_1 \frac{J}{2} I(\tau_J \geq J) | M_t = (n, s)\right] \tag{36}$$

where $\delta_1 = \min\{1, (1 - 3\gamma)/3\gamma\}$ has been defined in (22). We show that the first two terms on the right-hand side of (36) are bounded. Since (33) $\lim_{J\to\infty} -\delta_1 J/2 P[\tau_J \geq J] = -\infty$, this will be sufficient to prove that $\lim_{J\to\infty} E[(V(M_{t+J}) - V(M_t)) I(\tau_J \geq J) | M_t = (n, s)] = -\infty$. Define $W_k = X_{t+k} - X_t + k\gamma_1/2$ and $\tilde{F}_k = F_{t+k}$, where F_t is the sigma-field generated by $\{A_s, s \leq t - 1; X_s, s \leq t\}$, representing the history of the process $(M_t)_{t\geq 0}$ up to time t. To prove that the first term in (36) is bounded, we show that there exists $\phi > 0$ such that $(Y_k, \tilde{F}_k)$ is a supermartingale, with $Y_k = e^{\phi W_k} I(\tau_J \geq k)$. We need to show that $E[Y_{k+1} | \tilde{F}_k] \leq Y_k$, which is equivalent to

$$E[e^{\phi(X_{t+k+1} - X_t + (k+1/2)\delta_1)} I(\tau_J \geq k+1) | F_{t+k}] \leq e^{\phi(X_{t+k} - X_t + (k/2)\delta_1)} I(\tau_J \geq k)$$

since $I(\tau_J \geq k+1) = I(\tau_J \geq k) I(T_J \geq k+1)$, and $I(\tau_J \geq k)$ is measurable with respect to F_{t+k}

$$I(\tau_J \geq k) E[e^{\phi(X_{t+k+1} - X_{t+k} + \delta_1/2)} | F_{t+k}] \leq I(\tau_J \geq k). \tag{37}$$

Now if $\tau_J \geq k$, then from (30), $M_{t+k} \in C(\gamma, 5\gamma) \cap U_M$. Lemma 2.2 in [11] states that if X is a random variable such that $|X|$ is stochastically dominated by an exponential type random variable Z, and if the expectation of X is strictly negative, $E|X| < -\epsilon$, then there exist two constants $\eta > 0$ and $\rho < 1$ such that $E[e^{\eta X}] < \rho < 1$. Hence, there exists $\phi > 0$ such that

for all $(n, s) \in C(-5\gamma, 5\gamma) \cap U_M$,

$$E[e^{\phi(X_{t+1} - X_t + \delta/2)} | M_t = (n, s)] < 1 \tag{38a}$$

for all $(n, s) \in C(-\infty, -\gamma) \cap U_M$,

$$E[e^{\phi(\tilde{X}_{t+1} - \tilde{X}_t + \delta/2)} | M_t = (n, s)] < 1 \tag{38b}$$

for all $(n, s) \in C(\gamma, \infty) \cap U_M$,

$$E[e^{\phi(-\tilde{X}_{t+1} + \tilde{X}_t + \delta/2)} | M_t = (n, s)] < 1. \tag{38c}$$

It follows from (37) and (38a) that $(Y_k, \tilde{F}_k)$ is a supermartingale. Therefore,

$$E[Y_J | F_0] = E[e^{\phi W_J} I(\tau_J \geq J) | F_t] \leq E[Y_0 | \tilde{F}_0] = 1. \tag{39}$$

Finally, considering that $\max\{0, x\} \leq 1/\phi\, e^{\phi x}$, it follows from (39) that the first term in (36) is bounded. Using (30) and (38c), it can be shown with the same method that the second term in (36) is also bounded. Thus, there exists a constant B_T independent of J

uen that

$$E[(V(M_{t+J}-V(M_t))I(\tau_J\geq J)|M_t=(n,s)]\leq B_T-\tfrac{\nu}{2}\delta_1 P[\tau_J\geq J].$$

he case $(n,s)\in C(-4\gamma-\gamma/2,-2\gamma+\gamma/2)\cap U_{Q^*(J)}$ can be ealt with in a similar way, using (38a) and (38b). Therefore, we ave shown that there exist $\mu_2>0$ and $J_2>0$ such that for all J J_2 and for all $(n,s)\in Y_{Q^*(J)}\cap Z_P$

$$E[(V(M_{t+J})-V(M_t))I(\tau_J\geq J)|M_t=(n,s)]<-J\mu_2. \quad (40)$$

t is shown in [9] that there exist a constant $B>0$, a function $\nu_3(J)$ with $\lim_{J\to\infty}\nu_3(J)$, and a nonnegative function $\nu_J(M)$ epending on J verifying $\lim_{M\to\infty}\nu_J(M)=0$, such that for all $(n,$ $)\in U_{Q^*(J)+M_1}\cap Z_P$,

$$E[(V(M_{t+J})-V(M_t))I(\tau_J<J)|M_t=(n,s)]$$
$$B+\nu_1(J)+\nu_J(M_1). \quad (41)$$

Ve are now ready to conclude the proof of Proposition 3. From 40) and (41), we have for all $(n,s)\in U_{Q^*(J)+M_1}\cap Z_P$, $E[V(M_{t+J})-V(M_t)|M_t=(n,s)]\leq B-J\mu_2+\nu_3(J)+\nu_J(M_1)$. Fix an integer $J_f\geq\max\{J_1,J_2\}$ such that for all $J\geq J_f$, $B-J\mu_2+\nu\nu_3(J)<-\mu_2$. Then for all $(n,s)\in U_{Q^*(J_f)+M_1}\cap$ Z_P, we have $E[V(M_{t+J_f})-V(M_t)|M_t=(n,s)]\leq-\mu_2+\nu_J(M_1)$. On the other hand, we also have from (44), for all (n,s) $\in U_{Q^*(J_f)+M_1}\cap Z_P$

$$E[V(M_{t+J_f})-V(M_t)|M_t=(n,s)]\leq-\mu_1 J_f.$$

Now fix M_1 large enough so that $\nu_J(M_1)\leq\mu_2/2$. Then define M_f $=J'(J_f)+M_1$, and $\rho=\min\{\mu_2/2,J_f\mu_1\}$. □

We can now conclude that $(M_t)_{t\geq0}$ is geometrically ergodic for $\lambda<\eta_c$ by invoking the following result.

Theorem (Hajek [11]): Let $\{W_t\}$ be a sequence of random variables adapted to an increasing family of σ-fields $\{F_t\}$. Suppose that W_0 is deterministic, that $\{W_t,F_t\}$ is exponential type, and that for some $\epsilon>0$ and $a>0$ we have $E[(W_{t+1}-W_t)\cdot I(W_t>a)|F_t]\leq0$ for all $t\geq0$. Then for each value of W_0 the stopping time $\tau=\min\{t\geq0;W_t\leq a\}$ is exponential type.

Define $W_t=V(M_{tJ_f})$ and $a=M_f\max\{1,(1+3\gamma)/3\gamma,(1-\gamma)/3\gamma\}$. If $V(M_t)>a$, then $M_t\in U_{M_f}$. From (24) and C0 $V(M_t),F_t)$ is exponential type since A_t is. From Proposition 3, we can apply Hajek's result to our system to conclude that $\tau=$ $\min\{t\geq0,V(M_{tJ_f})\leq a\}$ is exponential type for any initial state. Since $V(M_t)\leq a$ implies that $X_t\leq a$ and $S_t\leq a/(1-\gamma)$, it follows that $\tau'=\min\{t\geq0,X_{tJ_f}\leq a$, and $S_{tJ_f}\leq a/(1-\gamma)\}$ is also exponential type for any initial state, as well as τ'' $=\min\{t\geq0,X_t\leq a$, and $S_t\leq a/(1-3\gamma)\}$. Hence, it follows from [14] that (X_t,S_t) is geometrically ergodic, concluding the proof of Theorem 4. □

V. Stability Proof Via Mikhailov's Theorem

Mikhailov [19, Theorem 3] has recently found a powerful sufficient condition to guarantee the stability of a Markov process taking values on $R^+\times R^+$. This result can be used to weaken the sufficient conditions we imposed in Section III and obtain a much more simple proof of stability. However, the form of stability used by Mikhailov is weaker than the geometric ergodicity used in Section III.

Let M_t be a discrete-time Markov process taking values in $Y\subseteq R^n$. $U(r)=\{x\in R^n:\|x\|\leq r\}$, and $\tau_x(S)=\min\{t\geq0:M_t\in S|M_0=x\}$, i.e., $\tau_x(S)$ is the time it takes to reach the set S from x. Then we say the process M_t is stable if there exist constants c_1 and c_2 such that $E[\tau_x(U(r))]\leq c_1\|x\|+c_2$ for all $x\in Y$. Using this definition of stability we show the following result which is analogous to Theorem 4.

Theorem 5: Suppose that:

i) the number of new packet arrivals per slot has finite second moment $E[A_t^2]<+\infty$;

ii) there exists $A\in(0,+\infty)$ such that $t(A)=\sup_{x\geq0}t(x)$;

iii) C0″: there exists $B<+\infty$ such that for all $n\geq1$, $\Sigma_{k=1}^n k^2\epsilon_{nk}\leq B$.

Fix $\lambda<\eta_c$ and $\xi>0$ such that $\lambda<t(A\xi)$. Choose $\alpha<0$ and $\beta>0$ such that

$$C1'\colon\ \beta(e^{A\xi}-1)=\frac{\lambda-t(A\xi)}{\xi}e^{A\xi}-\alpha$$

$$C_2'\colon\ \beta>m_\xi(\lambda)=\sup_{x>0,x\neq\xi}\frac{\lambda-t(Ax)-xe^{-A(x-\xi)}\dfrac{\lambda-t(A\xi)}{\xi}}{x-xe^{-A(x-\xi)}}.$$

Then the control algorithm

$$\gamma_t=\frac{A}{S_t}$$

$$S_{t+1}=\max\{A,S_t+\alpha I(Z_t=0)+\beta I(Z_t=\bar{0})\}$$

is stable.

Proof of Theorem 5: Let us state first Mikhailov's Theorem (cf. [35] for an exposition of this result and its application in the decentralized control of the conventional collision channel).

Theorem (Mikhailov [19]): Let $M_t=(X_t,S_t)$ be a homogeneous Markov process on $R^+\times R_0^+$ with drifts

$$(c(n,s),e(n,s))=E[M_{t+1}-M_t|M_t=(n,s)].$$

Suppose that:

i) there exists $B<+\infty$ such that for all $(n,s)\in R^+\times R_0^+$, $E[\|M_{t+1}-M_t\|^2|M_t=(n,s)]\leq B$;

ii) for all $\psi\in(0,+\infty)$, the drifts $(c(n,n/\psi),e(n,n/\psi))$ converge uniformly in ψ as n goes to infinity to $(c(\psi),e(\psi))$;

iii) the limit drifts $(c(\psi),e(\psi))$ are differentiable on $[0,+\infty)$, with $(e(0),e(0))=\lim_{s\to-\infty}(c(0,s),e(0,s))$;

iv) there exists $\epsilon>0$ such that if $c(\psi_0)=\psi_0 e(\psi_0)$, then $c(\psi_0)<-\epsilon$.

Then M_t is stable.

Since both the new packet arrivals and the rows of the reception matrix have finite variance, it is easy to check that condition i) in Mikhailov's Theorem holds

$$E[\|M_{t+1}-M_t\|^2|M_t=(n,s)]$$
$$=E[(X_{t+1}-X_t)^2+(S_{t+1}-S_t)^2|M_t=(n,s)].$$

Now $E[(S_{t+1}-S_t)^2|M_t=(n,s)]\leq\alpha^2+\beta^2$, and from (2)

$$E[(X_{t+1}-X_t)^2|M_t=(n,s)]\leq E[A_t^2]+E[\Sigma_t^2|M_t=(n,s)].$$

From C0′ the variance of the number of successes is also bounded

$$E[\Sigma_t^2|M_t=(n,s)]$$
$$=\sum_{k=1}^n k^2\sum_{j=k}^n\binom{n}{j}\left(\frac{A}{s}\right)^j\left(1-\frac{A}{s}\right)^{n-j}\epsilon_{jk}\leq B.$$

It follows directly from (16) and (17) that the limit drifts are given by

$$c(\psi)=\lambda-t(A\psi)$$
$$e(\psi)=\beta+(\alpha-\beta)e^{-A\psi},$$

respectively, for $\psi\in[0,+\infty)$. Uniform convergence to the limit drifts follows immediately from the results given for the perfect state information case (Property 4). Also it is clear the $t(x)$ is

differentiable (see (6), where $0 \le C_n \le n$). Therefore, properties ii) and iii) in Mikhailov's Theorem are satisfied.

In order to check property iv) note that if $\psi_0 = \xi$, then it follows from C1′ that

$$c(\psi_0) = \psi_0 e(\psi_0).$$

But, at that point, $c(\psi_0) < 0$ because of the choice of ξ. There is no other root of the equation $c(\psi) = \psi e(\psi)$, and, therefore, property v) follows. To see this, note that because of C1′, $c(\psi) = \psi e(\psi)$ for $\psi \neq \xi$ is equivalent to

$$\beta = \frac{\dfrac{\lambda - t(A\psi)}{\psi} - e^{A(\xi-\psi)}\dfrac{\lambda - t(A\xi)}{\xi}}{1 - e^{A(\xi-\psi)}}$$

which is impossible if $\psi \neq \xi$ because of C2′. □

It can be shown [9] that $m_\xi(\lambda)$ is finite for all nonnegative λ and ξ, and therefore the set of control laws defined by C1′ and C2′ is nonempty. Actually, the set of control laws in Theorem 4 is a subset of those in Theorem 5 because in Theorem 5 we can choose $\xi = 1$, in which case C2 is equivalent to C1′ and C1 is more restrictive than C2′ because $\lambda \ge m_1(\lambda)$ [9]. □

V. Conclusion

In this paper we have investigated the properties of decentralized control algorithms for a random access channel with multipacket reception capability. By using the working hypothesis that the users are aware of the value of the backlog, we have determined the best throughput achievable by any such protocol, as well as a simple way to achieve it. The optimum throughput has been shown to be given by the maximum average number of successes per slot when the number of transmissions, per slot is Poisson distributed. In the imperfect state information case, we have shown that the same throughput achieved in the perfect state information case can be achieved by using in lieu of the true backlog, an estimate of the backlog computed at each station using binary feedback, and we have used this estimate to derive a control scheme which is optimal in the sense that it achieves the optimal throughput determined earlier. This is true provided the reception matrix verifies condition C0, which puts some restrictions on the number of successes per slot. By using Mikhailov's result, C0 can be replaced by the weaker condition C0′. In this case however, geometric ergodicity was not ensured. Note that the feedback empty/nonempty used in Sections III and IV may be less than the available feedback in many practical situations, but no further information is needed: a ternary feedback would not shorten the proof or achieve better throughput.

Finally, let us mention that one can easily modify the proof of Theorem 4 to show that a similar result holds with the IFT access rule. More precisely, under a hypothesis paralleling those of Theorem 4, one can build a control scheme based on a binary feedback empty/nonempty such that the Markov vector (X_t, S_t) is geometrically ergodic for $\lambda < \sup_{x \ge 0} T(x)$. Using Theorem 3, it can be seen that the maximum stable throughput is the same for both access rules when the new packet arrivals are Poisson distributed.

Appendix

Kaplan's Condition

Consider a Markov chain with denumerable state-space D, and one-step transition probability matrix $(P_{xy})_{(x,y)\in D}$. Let $V(x)$ by a Lyapunov function on D. Then the generalized Kaplan's condition holds if there exists a positive constant B such that for all $z \in [0, 1[$ and all $x \in D$

$$z^{V(x)} - \sum_{y \in D} P_{xy} z^{V(y)} \ge -B(1-z).$$

1) One-Dimensional Kaplan's Condition: Consider the model of Section II with a control scheme $p_n = F(X_n)$, and the Lyapunov function $V(x) = x$. To check Kaplan's condition, it is enough from [27] to show that the downward part of the drift $-D(i) = \sum_{k=1}^{i} kP_{i,i-k}$ is bounded below. For $i \ge 1$ and $1 \le k \le i$ we have

$$P_{i,i-k} = \sum_{n=0}^{i-k} \lambda_n \sum_{j=k+n}^{i} \binom{i}{j} F(i)^j (1-F(i))^{i-j} \epsilon_{j,k+n}.$$

After a change of variable, it follows that

$$D(i) = \sum_{j=1}^{i} \binom{i}{j} F(i)^j (1-F(i))^{i-j} \sum_{n=0}^{j-1} \lambda_n \sum_{k=n+1}^{j} (k-n)\epsilon_{j,k}. \tag{A-1}$$

If $(C_n)_{n\ge 1}$ is bounded, then Kaplan's condition holds independent of the retransmission policy. Denoting by B_c an upper bound for $(C_n)_{n\ge 1}$, (A-1) becomes

$$-D(i) \ge -\sum_{j=1}^{i} \binom{i}{j} F(i)^j (1-F(i))^{i-j} \sum_{n=0}^{j-1} \lambda_n C_j$$

$$\ge -\sum_{j=1}^{i} \binom{i}{j} F(i)^j (1-F(i))^{i-j} C_j \ge -B_c. \tag{A-2}$$

2) Two-Dimensional Kaplan's Condition: Consider now the multipacket channel with a general control algorithm (1). Then (X_t, S_t) is the Markov chain of interest, and the relevant Lyapunov function is $V(n, s) = n$. We prove again that Kaplan's condition holds provided that $(C_n)_{n\ge 1}$ is bounded. From [27], it is enough also in this case to show that the downward part $T(x)$ of the generalized drift is bounded below, with $T(x) = \sum_{y/V(y)<V(x)} P_{xy}(V(y) - V(x))$. Given a state $x = (i, s)$, we have

$$T(x) = -\sum_{r=1}^{i} r \sum_{k} P[X_{n+1} = i-r,\ S_{n+1} = k \mid X_n = i,\ S_n = s]$$

$$= -\sum_{r=1}^{i} rP[X_{n+1} = i-r \mid x_n = i,\ S_n = s]$$

which is, in the same way as before

$$T(x) = -\sum_{r=1}^{i-k} r \sum_{n=0}^{i-r} \lambda_n \sum_{j=r+n}^{i} \binom{i}{j} (F(s))^j (1-F(s))^{i-j} \epsilon_{j,r+n}$$

$$= -\sum_{j=1}^{i} \binom{i}{j} F(s)^j (1-F(s))^{i-j} \sum_{n=0}^{j-1} \lambda_n \sum_{r=n+1}^{j} (r-n)\epsilon_{jr}$$

this expression is similar to (A-1), and the end of the proof is the same as in (A-2).

References

[1] N. Abramson, "The throughput of packet broadcasting channels," *IEEE Trans. Commun.*, vol. COM-25, pp. 117–128, 1977.

[2] J. C. Arnbak and W. van Blitterswijk, "Capacity of slotted Aloha in Rayleigh fading channels," *IEEE J. Select. Areas Commun.*, vol. SAC-5, no. 2, pp. 261–269, 1987.

[3] D. H. Davis and S. A. Gronemeyer, "Performance of slotted Aloha random access with delay capture and randomized time of arrival," *IEEE Trans. Commun.*, vol. COM-28, pp. 703–710, 1980.

[4] G. Fayolle, E. Gelenbe, and J. Labetoulle, "Stability and optimal control of the packet switching broadcast channel," *J. Ass. Comput. Mach.*, vol. 24, no. 3, pp. 375–386, 1977.

[5] F. G. Foster, "On the stochastic matrices associated with certain queuing processes," *Ann. Math. Stat.*, no. 24, pp. 355–360, 1953.
[6] S. Ghez, S. Verdú, and S. C. Schwartz, "Stability of multipacket Aloha," in *Proc. 21st Conf. ISS*, The Johns Hopkins Univ., Baltimore, MD, Mar. 1987.
[7] ——, "On decentralized control algorithms for multipacket Aloha," in *Proc. 25th Allerton Conf. Commun., Contr., Comput.*, Oct. 1987.
[8] ——, "Stability properties of slotted Aloha with multipacket reception capability," *IEEE Trans. Automat. Contr.*, vol. 33, July 1988.
[9] S. Ghez, "Random access communications for the multipacket channel," Ph.D. dissertation, Dept. Elec. Eng., Princeton Univ., 1989.
[10] D. J. Goodman and A. A. M. Saleh, "The near/far effect in local Aloha radio communications," *IEEE Trans. Vehic. Technol.*, vol. VT-36, no. 1, Feb. 1987.
[11] B. Hajek, "Hitting-time and occupation-time bounds implied by drift analysis with applications," *Adv. Appl. Prob.*, vol. 14, pp. 502–525, 1982.
[12] B. Hajek and T. Van Loon, "Decentralized dynamic control of a multiaccess broadcast channel," *IEEE Trans. Automat. Contr.*, vol. AC-27, pp. 559–569, 1982.
[13] M. Kaplan, "A sufficient condition for nonergodicty of a Markov chain," *IEEE Trans. Inform. Theory*, vol. IT-25, pp. 470–471, 1979.
[14] J. G. Kemeny, J. L. Snell, and A. W. Knapp, *Denumerable Markov Chains*. New York: Springer-Verlag, 1976.
[15] F. P. Kelly, "Stochastic models of computer communication systems," *J. R. Statist. Soc. B*, vol. 47, no. 3, pp. 379–395, 1985.
[16] C. C. Lee, "Random signal levels for channel access in packet radio networks," *IEEE J. Select. Areas Commun.*, vol. SAC-5, no. 6, pp. 1026–1034, 1987.
[17] N. Mehravari, "Collision resolution in random access systems with multiple reception," preprint, 1987.
[18] J. J. Metzner, "On improving utilization in Aloha networks," *IEEE Trans. Commun.*, vol. COM-24, pp. 447–448, Apr. 1976.
[19] V. A. Mikhailov, "Geometrical analysis of the stability of Markov chains in R_+^n and its application to throughput evaluation of the adaptive random access algorithm," *Problemy Peredachi Informatsii*, vol. 24, no. 1, pp. 61–73, Jan.–Mar. 1988; translated in *Problems of Inform. Transmission*, July 1988.
[20] C. Namislo, "Analysis of mobile radio slotted Aloha networks," *IEEE J.Select. Areas Commun.*, vol. SAC-2, pp. 583–588, 1984.
[21] A. G. Pakes, "Some conditions for ergodicity and recurrence of Markov chains," *Operat. Res.*, vol. 17, pp. 1058–1061, 1969.
[22] A. Polydoros and J. Silvester, "Slotted random access spread-spectrum networks: An analytical framework," *IEEE J. Select. Areas Commun.*, vol. SAC-5, no. 6, pp. 989–1002, 1987.
[23] B. Ramamurthi, A. A. M. Saleh, and D. J. Goodman, "Aloha with perfect capture," preprint 1986.
[24] D. Raychaudhuri, "Performance analysis of random access packet-switched code division multiple access systems," *IEEE Trans. Commun.*, vol. COM-29, pp. 895–901, 1981.
[25] R. L. Rivest, "Network control by Bayesan broadcast," *IEEE Trans. Inform. Theory*, vol. IT-33, no. 3, pp. 323–328, 1987
[26] L. G. Roberts, "Aloha packet system with and withpout slots and capture," *Comput. Commun. Rev.*, no. 5, pp. 28–42, 1975.
[27] L. I. Sennott, "Tests for the nonergodicity of multidimensional Markov chains," *Operat. Res.*, vol. 33, pp. 161–167, 1985.
[28] A. Schwartz and M Sidi, "Erasure capture and noise errors in controlled multiple access networks," in *Proc. 25th Conf. Decision Conf.*, Dec. 1986, pp. 1333–1334.
[29] M. K. Simon, J. K. Omura, R. A. Scholtz, and B. K. Levitt, *Spread-Spectrum Communications*. New York: Computer Science Press, 1985.
[30] J. N. Tsitsiklis, "Analysis of a multiaccess control scheme," *IEEE Trans. Automat. Contr.*, vol. AC-32, pp. 1017–1020, 1987.
[31] B. S. Tsybakov, V. A. Mikhailov, and N. B. Likhanov, "Bounds for packet transmission rate in a random multiple access system," *Problemy Peredachi Informatsii*, vol. 19, no. 1, pp. 61–81, 1983
[32] R. L. Tweedie, "Sufficient conditions for ergodicity and recurrence of Markov chains on a general state space," *Stoch. Proc. Appl.*, no. 3, pp. 385–403, 1975.
[33] S. Verdú, "Minimum probability of error for asynchronous Gaussian multiple-access channels," *IEEE Trans. Inform. Theory*, vol IT-32, pp. 85–96, 1986.
[34] J. E. Wieselthier, A. Ephremides, and L. A. Michaels, "An exact analysis and performance evaluation of framed Aloha with capture," preprint, 1987.
[35] A. Ephremides and S. Verdú, "Control and optimization methods in communication network problems," *IEEE Trans. Automat. Contr.*, vol. 34, pp. 930–942, Sept. 1989.

Sylvie Ghez (S'80) was born in Paris, France, on June 3, 1961. She received the engineering diploma in 1984 from the Ecole Nationale Supérieure des Télécommunications, Paris, and the M.S. degree in electrical engineering from Princeton University, Princeton, NJ, in 1987.

At present she is a doctoral candidate at Princeton University. Her research interests include multiple access protocols, network routing problems, queueing theory, and applied probability theory

Ms. Ghez was the recipient of the Princeton Wallace Memorial Fellowship for the year 1987–1988.

Sergio Verdú (S'80–M'84–SM'88), for a photograph and biography, see p 942 of the September 1989 issue of this TRANSACTIONS.

Stuart C. Schwartz (S'64–M'66–SM'83) was born in Brooklyn, NY, on July 12, 1939. He received the B.S. and M.S. degrees in aeronautical engineering from the Massachusetts Institute of Technology, Cambridge, in 1961, and the Ph.D. degree from the Information and Control Engineering Program, University of Michigan, Ann Arbor, in 1966.

While at M.I.T. he was associated with the Naval Supersonic Laboratory and the Instrumentation Laboratory. During the year 1961–1962 he was at the Jet Propulsion Laboratory, Pasadena, CA, working on problems in orbit estimation and telemetry He is presently a Professor of Electrical Engineering and Chairman of the department at Princeton University, Princeton, NJ. He served as Associate Dean of the School of Engineering during the period July 1977–June 1980 During the academic year 1972–1973 he was a John S. Guggenheim Fellow and a Visiting Associate Professor in the Department of Electrical Engineering, Technion—Israel Institute of Technology, Haifa, Israel During the academic year 1980–1981, he was a member of the Technical Staff at the Radio Research Laboratory, Bell Telephone Laboratories, Crawford Hill, NJ His principal research interests are in the application of probability and stochastic processes to problems in statistical communication and system theory

Dr. Schwartz is a member of Sigma Gamma Tau, Eta Kappa Nu, and Sigma Xi. He has served as an Editor for *SIAM Journal on Applied Mathematics* and as Program Chairman for the 1986 ISIT.

An Exact Analysis of the Optical CDMA Noncoherent Receiver

David Brady, Sergio Verdú †
Department of Electrical Engineering
Princeton University,
Princeton, NJ 08544

Abstract

This work studies optical Code Division Multiple Access (CDMA) systems, and presents the exact error expression for the noncoherent, single-user matched-filter receiver based on the electron count in a symbol period. This analysis is valid for arbitrary photomultipliers, adheres fully to the semi-classical model of light, and does not depend on approximations for large user groups or strong received optical fields

The general error rate expression is specialized to the case of unity gain photodetectors and prime sequences, and the exact minimum probability of error and optimal threshold are compared to those obtained with simplifying assumptions on user transmission coordination or multiple-access-interference (MAI) distribution. We find that the approximation of chip synchronism yields a weak upper bound on the true error rate, and we demonstrate that the approximations of perfect optical-to-electrical conversion and Gaussian MAI yield an optimal hypothesis test whose error rate overestimates the true minimum error rate and underestimates the optimal threshold for moderate and large received optical energies.

Optical CDMA Model

The digital modulation format studied in this paper is optical Direct Sequence Spread Spectrum, i.e., during each symbol interval of duration T, the j^{th} transmitting laser is amplitude- modulated by the product of the data, which takes on values in $\{0,1\}$, and an assigned, signature sequence of relatively short rectangular pulses. This scheme divides the symbol interval into N equal length subintervals, called chips, on which the signature sequence is constant and takes on values in $\{0,1\}$. Further, we define $P_j = P$ as the number of non-zero chips in each signature sequence, $b_{j,n}$ as the transmitted symbol of the j^{th} user in the interval $[nT,(n+1)T]$, and $c_j(t)$ as a periodic replication of the signature sequence of the j^{th} user such that $c_j(t)$, $t \in [nT,(n+1)T)\}$ is the j^{th} signature sequence for any fixed integer n. Then the transmitted complex scalar field from the j^{th} laser may be expressed as

$$r_j(t) = \sqrt{\frac{s\,N}{T}} c_j(t) b_{jn} e^{j(\nu t + a_j W_j(t) + \theta_j)}, \qquad nT \le t - \tau_j < (n+1)T \tag{1}$$

where s is proportional to the optical energy per bit of the transmitting laser, ν denotes the optical carrier frequency (assumed to be identical for all users), and θ_j is the phase offset of the j^{th} laser from the first laser. In this expression $W_j(t)$ is a standard Brownian motion, and a_j is related to the j^{th} transmitting laser linewidth, B_j, by $a_j = \sqrt{2\pi B_j}$. The relative delays $\{\tau_j\}$ are defined on $[0,T)$ with reference to the receiver of the first user. With dispersion-free transmission (1) also represents the complex scalar field at the first receiver due to user j.

We shall assume that the symbol rate of each user is the same, the optical fields of the K users add in a noncoherent fashion, and that each single-user receiver acquires the timing of its transmitter's symbol epochs. As there is no cooperation between the users, it is appropriate to model the remaining relative delays, $\{\tau_j\}_{j=2}^{K}$, as independent, identically distributed, random variables that are uniformly distributed on the interval $[0,T]$. It follows that the intensity of the optical field at the receiver of the first user is

$$|r(t)|^2 = \frac{sN}{T} \sum_{j=1}^{K} b_{j,-1} c_j(t-\tau_j) p_l(0,\tau_j) + b_{j,0} c_j(t-\tau_j) p_l(\tau_j,T)$$

Where $p_l(a,b)$ a rectangular pulse of unit height with support $[a,b)$. Due to the modulation shown in (1), the resulting photon point process depends on the data $b_{1,0}$ only on the set $\{t|c_1(t) = 1,\ 0 \le t < T\}$. A commonly used receiver for this channel is the noncoherent matched-filter, which sums the photon counts in each of the nonzero chip subintervals of the user of interest. Given that the function $c_1(t)$ takes values on $\{0,1\}$, the correlation operation would be easily achieved at extremely low chip rates by an electro-optic modulator, which would allow received light to pass only when $c_1(t) = 1$. A more effective device to achieve the matched-filtering operation at higher chip rates is the fiber optic tap delay line, which uses the finite

This work was partially supported by the U S Army Research Office under Contract DAAL03-87-K-0062 and by the U S. Office of Naval Research under Contract N0014-87-K-0054

To appear in the *Proceedings of the Conference on Information Sciences and Systems*, Johns Hopkins University, 1989

ropagation velocity of light to achieve the proper rel ative delay of two optical signals by passing them through ibers of different lengths. The matched-filter direct-letection receiver has been studied in several experiments 1,2] and will be the CDMA receiver analyzed in this work

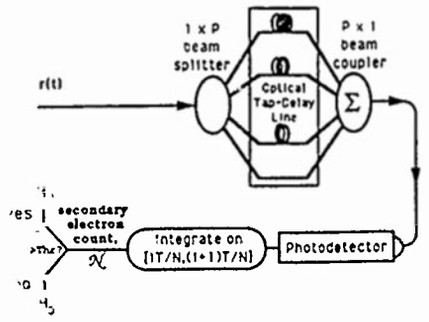

Figure 1. Optical Noncoherent Matched-Filter CDMA Receiver

as shown in Figure 1, the total received optical signal rt) is coupled to a $1 \times P$ beam splitter. Each of the outputs n the splitter are identical copies of the input signal, only ttenuated in intensity by P. These signals are input to the ap delay line. The function of the i^{th} tap is to delay the eceived field so that the optical signal in the i^{th} non-zero hip of the first signature sequence overlaps in time with he last (P^{th}) non-zero chip of the same bit interval in the naelayed signal. Thus, the first tap requires more fiber ayle than the second. The tapped signals are noncoher-ntly recombined, and the output optical signal is incident on the photodetector. To decide on the value of $b_{1,0}$, we se the secondary electron count during the last non-zero hip interval of the first signature sequence For the re-nainder of this work we denote this secondary electron ount by $\mathcal{N}$ We shall employ a common photomultiplier nodel, in which the intensity of primary electrons is given by $\alpha|r(t)|^2 + \beta$, where α is proportional to the quantum efficiency of the photodetector, and β denotes the rate of primary electrons due to an independent dark current The n^{th} primary electron yields a random number of secondary (output) electrons g_n, and the collection $\{g_n\}$ is assumed to be mutually independent, identically distributed, and independent of the photon or primary electron point process [3]. The common probability generating function of $\{g_n\}$ is denoted as $G(z) = \sum_{k=0}^{\infty} p_k z^k$. In this case, $\mathcal{N}$ is conditionally compound Poisson given the integrated intensity, which we define as Λ, and the distribution of $\mathcal{N}$ depends only on $G(z)$ and the integrated intensity Λ, given by

$$\Lambda \triangleq \alpha s\, b_{1,0} + d + \frac{\alpha s}{P} \sum_{j=2}^{K} b_{j,-1} R_{j,1}(\tau_j) + b_{j,0} \dot{R}_{j,1}(\tau_j) \quad (2)$$

where $R_{j,1}(\tau)$ and $\dot{R}_{j,1}(\tau)$ are the normalized (partial) cross correlations

$$R_{j,1}(\tau) \triangleq \frac{N}{T} \int_0^{\tau} c_j(t-\tau) c_1(t)\, dt$$

$$\dot{R}_{j,1}(\tau) \triangleq \frac{N}{T} \int_{\tau}^{T} c_j(t-\tau) c_1(t)\, dt$$

hat represent the contributions to the conditional mean Λ ue to the j^{th} signature sequence for the duration of $b_{j,-1}$ and $b_{j,0}$, respectively. Also, d represents the portion of the primary electron count mean due to thermoelectrons. In he remainder of this work we set the quantum efficiency of the photodetector to unity, as this effects the distribution of $\mathcal{N}$ only through an attenuation of intensity. Further, we set $x \triangleq \alpha s b_{1,0} = 0$ under hypothesis $\mathcal{H}_0$ and $x = s$ under hypothesis $\mathcal{H}_1$.

Derivation of $\mathcal{P}\left[\mathcal{N} = n \,\middle|\, x\right]$

In this section we obtain the general expression for the PMF of the secondary electron count $\mathcal{N}$, at the integrator output for an arbitrary photomultiplier and for synchronous or asynchronous transmission. We will use this result in a later section to compare the error rates under various simplifying approximations to the exact error rate. Also, the form of the general expression will be used in the next section to develop arbitrarily tight, computationally efficient bounds on the cumulative distribution function of $\mathcal{N}$.

In the following, we define M as the upper bound on the set of total cross-correlations $R_{j,k} + \dot{R}_{j,k}$, and as the signature sequences are from $\{0,1\}^N$, these bounds hold for the partial cross-correlations as well. Since the relative delays are uniformly distributed and the chip waveform is rectangular, it is straightforward to show that each cross-correlation is a mixed random variable whose measures have point masses on the integers $\{0,1,\ldots,M\}$ and continuous portions that are constant between these integers We shall employ the following notation

$$\mathcal{P}\left[R_{j,1} = i\right] = d_j(i) \qquad i \in \{0,1,\ldots,M\}$$

$$\mathcal{P}\left[R_{j,1} \in (v, v+dv)\right] = c_j(i)\, dv \qquad (v, v+dv) \in (i, i+1)$$

and we denote the distribution of $R_{j,1}$ as $\{d_j(0), d_j(1), \ldots, d_j(M), c_j(0), \ldots, c_j(M-1)\}$. Thus the marginal distribution of each cross-correlation is completely specified by $2M$ parameters. Further, the superscript-T notation will be used to distinguish the distribution of the total cross-correlation $R_{j,1} + \dot{R}_{j,1}$ from that of $R_{j,1}$, and the hat notation will be used for the distribution of $\dot{R}_{j,1}$.

Our approach to finding the PMF of $\mathcal{N}$ is the following: we will derive the z-transform of $\mathcal{N}$ from its conditional compound Poisson nature, and then show that this z-transform has a particularly straightforward and explicit Maclaurin series expansion. The PMF is the collection of coefficients of this series, and may be explicitly represented.

By conditioning on $(x, \{R_{j,1}, \hat{R}_{j,1}\}, j = 2, \ldots K)$, the count $\mathcal{N}$ has a compound Poisson distribution, whose z-transform is given by

$$\mathrm{E}\left[z^{\mathcal{N}} \,\middle|\, x, \{R_{j,1}, \hat{R}_{j,1}\}, j = 2, \ldots K\right] = \tag{3}$$
$$e^{(x+d)(G(z)-1)} \times \prod_{j=2}^{K} e^{\frac{s}{P}(b_{j,-1}R_{j,1} + b_{j,0}\hat{R}_{j,1})(G(z)-1)}$$

Due to the mutual independence of the pairs $\{R_{j,1}, \hat{R}_{j,1}\}$ we need to determine only the expectation of each factor in (3), as the j^{th} factor depends only on the random mixture $b_{j,-1}R_{j,1} + b_{j,0}\hat{R}_{j,1}$. It is clear that the random mixture has the same kind of distribution as $R_{j,1}$, and we denote this mixed distribution by $(D_j(0), D_j(1), \ldots, D_j(M), C_j(0), \ldots, C_j(M-1))$. With this notation, the closed form expression of the power series of interest is

$$\mathrm{E}\left[z^{\mathcal{N}} \,\middle|\, x\right] = e^{(G(z)-1)(x+d)} \times \tag{4}$$
$$\prod_{j=2}^{K} \left\{ \sum_{q=0}^{M} D_j(q) \exp\left(q\frac{s}{P}(G(z)-1)\right) - \right.$$
$$\left. \frac{P}{s}\frac{e^{(G(z)-1)\frac{s}{P}} - 1}{1 - G(z)} \sum_{r=0}^{M-1} C_j(r) \exp\left(r(G(z)-1)\frac{s}{P}\right) \right\}$$

We are interested in finding $\mathcal{P}\left[\mathcal{N} = n \,\middle|\, x\right]$, which is the coefficient of z^n in the power series of (4) about the origin. This power series is straightforward but unnecessarily general for most signature sequence sets of interest. For example, the number of parameters in the power series is reduced by a factor of $K-1$ by assuming that the marginals of $R_{j,1}$, $R_{j,1}$ and $R_{j,1} + \hat{R}_{j,1}$ are independent of j, i.e. the contribution of user j to the MAI is statistically indistinguishable from the other interferers. We have verified that this is an excellent approximation when the signature sequences come from the prime codes, and will drop the subscript from the distribution of the random mixtures in the sequel. Also, the power series of this expression is concisely written if we define $C(-1) = C(M) = 0$. With these simplifications, (4) becomes

$$\mathrm{E}\left[z^{\mathcal{N}} \,\middle|\, x\right] = e^{(G(z)-1)(x+d)} \times \tag{5}$$
$$\left\{ \sum_{q=0}^{M} D(q) e^{q\frac{s}{P}(G(z)-1)} - \right.$$
$$\left. \frac{P}{s}\frac{1}{1 - G(z)} \sum_{q=0}^{M} [C(q-1) - C(q)] e^{q(G(z)-1)\frac{s}{P}} \right\}^{K-1}$$

There are 2M+2 terms inside of the braces. Letting n_q index the number of occurrences of $D(q)$, and m_q the number of occurences of $[C(q-1) - C(q)]$ in a multinomial expansion, we rewrite (5) as

$$\mathrm{E}\left[z^{\mathcal{N}} \,\middle|\, x\right] = \sum \frac{(K-1)!}{\prod_{q=0}^{M} n_q! m_q!} \prod_{q=0}^{M} D(q)^{n_q} \left[\frac{P}{s}\{C(q-1) - C(q)\}\right]^{m_q}$$
$$\times \frac{\exp[(G(z)-1)\{x + d + \frac{s}{P}\sum_{q=0}^{M} q[n_q + m_q]\}]}{(1 - G(z))^{\sum_{q=0}^{M} m_q}} \tag{6}$$

where the outer summation is over all the indices such that $\sum_{q=0}^{M} m_q + n_q = K - 1$. We find the PMF of $\mathcal{N}$ in the following way. Suppose that we knew explicitly the coefficients of the following power series

$$\sum_{n=0}^{\infty} \mathcal{R}es(n, \alpha, \beta) z^n \triangleq \frac{e^{\alpha(G(z)-1)}}{(1 - G(z))^{\beta}}, \quad \alpha \in \mathbb{R}, \tag{7}$$

Using (7) in (6) we express the PMF for $\mathcal{N}$ as

$$\mathcal{P}\left[\mathcal{N} = n \,\middle|\, x\right] = \sum \frac{(K-1)!}{\prod_{q=0}^{M} n_q! m_q!} \prod_{q=0}^{M} D_q^{n_q} \left[\frac{P}{s}\{C(q-1) - C(q)\}\right]^{m_q}$$
$$\mathcal{R}es\left(n, \{x + d + \frac{s}{P}\sum_{q=0}^{M} q[n_q + m_q]\}, \sum_{q=0}^{M} m_q\right) \tag{8}$$

All that remains to be determined is an explicit expression for the coefficients $\mathcal{R}es()$ of the power series in (7). In the following we show that $\mathcal{R}es$ may be calculated by a linear recursion on the integers n and β.

The recursion for $\mathcal{R}es$ is most easily seen by substituting the identity

$$\frac{e^{\alpha(G(z)-1)}}{(1 - G(z))^{\beta+1}} = G(z)\frac{e^{\alpha(G(z)-1)}}{(1 - G(z))^{\beta+1}} + \frac{e^{\alpha(G(z)-1)}}{(1 - G(z))^{\beta}}$$

where $\beta \in \{0, 1, 2, \ldots\}$, into the definition for $\mathcal{R}es$ (7). This yields

$$(1 - p_0)\mathcal{R}es(n+1, \alpha, \beta+1) = \sum_{l=1}^{n+1} p_l \mathcal{R}es(n+1-l, \alpha, \beta+1)$$
$$+ \mathcal{R}es(n+1, \alpha, \beta), \quad n, \beta \in \{0, 1, 2, \ldots\} \tag{9}$$

where $G(z) = \sum_{l=0}^{\infty} p_l z^l$. For most photomultiplier models $p_0 = 0$, which we will assume in the sequel. The initial conditions of this recursion are also easily extracted from the definition of $\mathcal{R}es$,

$$\mathcal{R}es(0, \alpha, \beta) = e^{-\alpha}, \quad \beta \in \{0, 1, 2, \ldots\} \tag{10}$$

$$\mathcal{R}es(n, \alpha, 0) = \sum_{k=0}^{n} \frac{\alpha^k}{k!} e^{-\alpha} \mathcal{P}\left[\sum_{l=1}^{k} g_l = n\right], \quad n \in \{0, 1, \ldots\}.$$

The linear recursion for $\mathcal{R}es$ on n and β permits fast, efficient computation for any arguments $n, \beta \geq 1$. Note

that the second initial condition for this recursion depends on $\mathcal{P}\left[\sum_{l=1}^{k} g_l = n\right]$, which must be known for $n,k \in \{0,1,2,\ldots\}$. These probabilities require iterated convolutions of the PMF of the random gain g_l, may be precomputed and stored for small n and k, and may be accurately approximated online for large n,k. We are naturally interested in special cases where $\mathcal{P}\left[\sum_{l=1}^{k} g_l = n\right]$ has an explicit form - it is easy to show that this is the case for random gains that are shifted Poisson-distributed, as well as for the unity gain case.

Arbitrarily Tight Bounds on $\mathcal{P}\left[\mathcal{N} \le n \,\middle|\, x\right]$

Computationally efficient bounds must reduce the complexity of (8) in both the multinomial summation and the computation of $\mathcal{R}es$, while controlling the loss of accuracy by a parameter of our selection. In this section we show that by quantizing the random mixtures, we achieve all three objectives.

The complexity of the PMF is due to the smoothing over the joint distribution of the random mixtures we originally conditioned on these random variables to take advantage of the conditional compound Poisson nature of $\mathcal{N}$. We could have also conditioned according to the conditional mean, Λ, for which $\mathcal{N}$ is also compound Poisson However, the exact distribution of the conditional mean Λ is not easily obtained, as it is formed by the convolution of $K-1$ mixed distributions. It is obvious that if the convolved distributions were discrete, say, with $QM+1$ points, then the exact distribution of $\mathcal{N}$ would be straightforward to compute. More importantly, the distribution of Λ would take on $(K-1)QM+1$ points, rather than a number that is exponential in the number of interferers.

But how do we obtain bounds on $\mathcal{P}\left[\mathcal{N} \le n \,\middle|\, x\right]$ that use a discrete distribution on Λ, and are arbitrarily tight? Suppose we quantize the random mixtures $\{b_{j,-1}R_{j,1} + b_{j,0}R_{j,1}\}$ with a $\frac{1}{Q}$ quantization step size, $Q \in \{1,2,\ldots\}$, and round-up or round-down to form bounds on the random mixtures That is, we form Λ_l, Λ_u given by

$$\Lambda_l = x + d + \frac{s}{P}\sum_{j=2}^{K} b_{j,-1}\frac{1}{Q}\lfloor QR_{j,1}\rfloor + b_{j,0}\frac{1}{Q}\lfloor QR_{j,1}\rfloor$$

$$\Lambda_u = x + d + \frac{s}{P}\sum_{j=2}^{K} b_{j,-1}\frac{1}{Q}\lceil QR_{j,1}\rceil + b_{j,0}\frac{1}{Q}\lceil QR_{j,1}\rceil$$

where $\lfloor R\rfloor$ ($\lceil R\rceil$) is the greatest (least) integer function of R. Then it is obvious that $\Lambda_l \le \Lambda \le \Lambda_u$, but can we use Λ_u, Λ_l to form bounds on the secondary electron count CDF?

A subtle point is raised by considering the form of $\mathcal{N}$

$$\mathcal{N}(\Lambda) = \sum_{p=1}^{\Pi(\Lambda)} g_p$$

where $\Pi(\Lambda)$ is the conditionally Poisson number of primary electrons with conditional mean Λ. Since g_p are non-negative, we have that $\mathcal{N}$ is an increasing function of the primary electron count, Π. It is not clear that (a.s.) bounds on Λ produce similar bounds on $\Pi(\Lambda)$, as $\mathcal{P}\left[\Pi(\Lambda_l) > \Pi(\Lambda) \,\middle|\, x\right] > 0$, and this representation of $\mathcal{N}$ does not guarantee bounds on $\mathcal{P}\left[\mathcal{N} \le n \,\middle|\, x\right]$. In the lemma below we use a statistically equivalent representation of $\mathcal{N}$ to show that we may achieve bounds on $\mathcal{P}\left[\mathcal{N} \le n \,\middle|\, x\right]$ by using the distributions of Λ_l, Λ_u.

Lemma. Let $\Pi(\Lambda)$ be a conditional Poisson random variable with mean Λ given Λ, and let $\mathcal{N}(\Lambda) = \sum_{k=1}^{\Pi(\Lambda)} g_k$, where $\{g_k\}$ are independent, identically distributed, non-negative integer-valued random variables. Let $\Lambda' < \Lambda$, a.s. Then

$$\mathcal{P}\left[\mathcal{N}(\Lambda) \le n\right] \le \mathcal{P}\left[\mathcal{N}(\Lambda') \le n\right], \quad n \ge 0.$$

Proof. We recall that $p_k = \mathcal{P}\left[g_j = k\right]$, and define $\{\mathcal{M}_k(\Lambda' p_k),\ \Pi_k(\Lambda p_k)\}$ to be a set of conditionally mutually-independent, Poisson random variables with the indicated means given (Λ, Λ') so that $\Pi(\Lambda) = \sum_{k=1}^{\infty} \Pi_k(\Lambda p_k)$ Under this conditioning, $\mathcal{N}(\Lambda)$ has the same distribution as [4]

$$\mathcal{N}(\Lambda) = \sum_{k=1}^{\infty} k\Pi_k(\Lambda p_k)$$

It is straightforward to show that if $\{X_1, X_2, Y_1, Y_2\}$ are conditionally mutually-independent random variables given Λ, Λ' and

$$\mathcal{P}\left[X_i \le n \,\middle|\, \Lambda, \Lambda'\right] \le \mathcal{P}\left[Y_i \le n \,\middle|\, \Lambda, \Lambda'\right], \quad i = 1,2$$

then the same is true for the sum

$$\mathcal{P}\left[X_1 + X_2 \le n \,\middle|\, \Lambda, \Lambda'\right] \le \mathcal{P}\left[Y_1 + Y_2 \le n \,\middle|\, \Lambda, \Lambda'\right].$$

Since the Poisson CDF is a decreasing function of the mean, we have for $I = 1$

$$\mathcal{P}\left[\sum_{k=1}^{I} k\Pi_k(\Lambda p_k) \le n \,\middle|\, \Lambda, \Lambda'\right] \le \mathcal{P}\left[\sum_{k=1}^{I} k\mathcal{M}_k(\Lambda' p_k) \le n \,\middle|\, \Lambda, \Lambda'\right]$$

The same is true for the unconditioned CDFs by smoothing. The same holds for finite I by induction on the above fact, and for $I \to \infty$ by monotone sequential continuity of the probability measure. ∎

Example: Prime Sequences and PIN Photodiodes

A necessary prerequisite to the comparison between error rates of the CDMA matched-filter receiver is the computation of the random mixture distribution $(D(0),\ldots, D(M), C(0),\ldots,C(M-1))$, as seen in (8). These are computed by a knowledge of the signature sequences, as well as the distribution of the relative delay. Since the cross-correlations of prime sequences are bounded above by [5] $M = 2$, we must compute $(D(0), D(1), D(2), C(0), C(1))$ for the chip-synchronous, and asynchronous cases. For the prime sequences from GF(31), we have found that the average distributions for the random mixtures are

$$\begin{aligned} & (D(0), D(1), D(2), C(0), C(1)) \\ \textit{chip synchronous} &\Rightarrow (.57, .36, .07, .00, .00) \\ \textit{asynchronous} &\Rightarrow (.44, .22, .01, .24, .09) \end{aligned}$$

As noted earlier, we have verified that the MAI for prime sequences is well-modeled by a sum of independent, identically distributed (IID) random variables in the sense that the mean, variance, and third central moment of the MAI using the IID assumption and the average distribution were identical to the exact MAI moments, while the fourth central moments differed by less than .004% for 29 interferers. Further, these distributions did not differ significantly for the prime sequences from GF(11) and GF(17), and we use these distributions for all calculations.

In Figure 2 we have plotted the minimum error probability of the matched-filter CDMA receiver for the chip-synchronous assumption and for completely asynchronous transmission. We have used the weight 17 and length 289 prime sequences from GF(17), a received optical energy of s=1000 photons per bit, and a dark current contribution of d=50 thermoelectrons per bit. For a bit rate of R bits per second, these numbers correspond to a peak received power of $R\ 10^{-7} mW$ and a photodetector dark current of approximately $R\ 10^{-8} nA$. From Figure 2 we see that in this particular case the chip synchronous approximation upper bounds the error rate in the asynchronous case by at least one order of magnitude.

The error rates are ordered in this way due exclusively to the differences of the distributions of the random mixtures shown above. Note that the means of the random mixtures are identical in both cases, while the ordering of the variances coincides with that of the error rates. Thus the MAI has identical means under these distributions, and second moments whose ordering coincides with that of the error rates. It is easy to show that $\mathcal{E}[\mathcal{N}\,|\,x] = \bar{\Lambda}$, and $Var(\mathcal{N}|x) = \bar{\Lambda} - (\bar{\Lambda})^2 + \overline{\Lambda^2}$, which implies that under each hypothesis on x the mean of $\mathcal{N}$ is unchanged by the approximation of chip synchronism, yet the variance of $\mathcal{N}$ given x increases as we proceed from complete asynchronism to chip-synchronism. From the ordering of the minimum error rate curves in Figure 2, we see that an increase in the variance of $\mathcal{N}$ under each hypothesis results in an increased error rate as the conditional means of $\mathcal{N}$ are fixed.

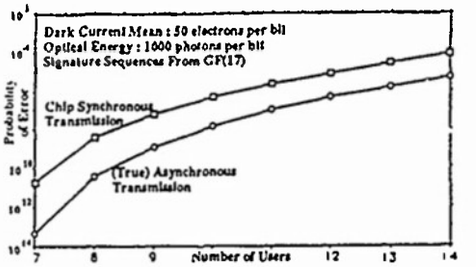

Figure 2. Comparison of the Minimum Error Rates For Complete Asynchronism and Chip Synchronous Approximation

Direct detection systems often require large received optical energies to achieve an acceptable error rate when a PIN photodiode is used, so we are interested in the asymptotic distribution of (a scaled version of) $\mathcal{N}$. The question is more formally worded as: if $\mathcal{N}$ is a conditionally-Poisson random variable with mean Λ given Λ, and $\frac{\Lambda-\bar{\Lambda}}{\sigma_\Lambda}$ tends in distribution to a random variable ϕ as some parameter grows without bound, what does the distribution of $\frac{\mathcal{N}-\bar{\mathcal{N}}}{\sigma_{\mathcal{N}}}$ tend to? In the simple case when Λ is deterministic, it is well known that the normalized count converges in distribution to a standard Gaussian random variable. Is this the case in general?

The answer was solved independently by Serfozo [6] and Grandell [7] for the special case when $\bar{\Lambda} \to \infty$, and depends on the limit ρ defined as $\lim \sigma_\Lambda^2/\bar{\Lambda}$. If $\rho = 0$, then the normalized count converges in distribution to a standard Gaussian. If $\rho = \infty$, then the normalized count converges in distribution to ϕ. Finally, if $0 < \rho < \infty$, then the normalized count converges in distribution to an independent mixture of a standard Gaussian and ϕ.

In our case, the parameter is the received signal energy per bit, s, and the condition $\bar{\Lambda} \to \infty$ is satisfied as Λ is proportional to s. It is this fact that also sets ρ to ∞, and we have from the result above that for large signal energies the normalized count converges in distribution to the scaled conditional mean ϕ. This asymptotic result is a weaker form of what is more commonly known as "perfect optical-to-electrical conversion", in which the integrated photocurrent is equal (a.s.) to the integrated optical intensity. It will be seen in the numerical results presented

next that the asymptotic statistic is far from being a deterministic signal in Gaussian noise, as the MAI is far from Gaussian even for a moderate number of users.

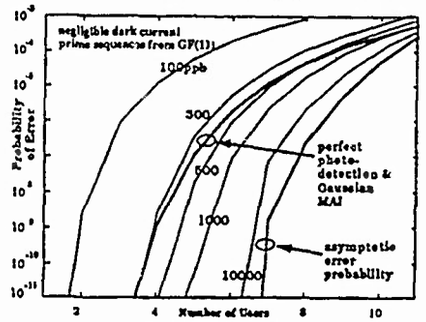

Figure 3. Exact Error Rates vs. Error Rates for Various Approximations

In Figure 3 we have compared the minimum error rates of the CDMA matched-filter receiver based on perfect optical-to-electrical conversion (the high energy limit) to those for the true distribution of $\mathcal{N}$ at various finite optical energies. In this example we have used the prime sequence from GF(11). Also, we have plotted the minimum error rate under the additional assumption of Gaussian-distributed MAI. We note that even for modest received optical energies of 10,000 photons per bit the error rate exceeds that predicted by the asymptotic distribution by at least an order of magnitude. Figure 3 shows that the minimum error rate is a decreasing function of the received optical energy, as expected. Further, we note that a Gaussian assumption on the MAI, together with the perfect optical-to-electrical assumptions is a poor estimate of the true minimum error rate curve, except for user group sizes exceeding, say, 10 users. In particular, this assumption overestimates the error rate for moderate to large incident optical energies.

As a result of the perfect optical-to-electrical-conversion approximation, the boundedness of the MAI leads to an "error-free" condition for sufficiently small numbers of interferers. This occurs since the supports of the conditional distributions of the test statistic are disjoint under these assumptions. Since prime sequences have cross-correlations that are bounded above by 2, the necessary condition for prime sequences is $K - 1 < P/2$. This assumption predicts zero error rate for $K \leq 6$ in Figure 3, which indicates that the perfect optical-to-electrical assumption accurately predicts the "error-free condition" only for incident optical energies exceeding 10,000 photons per bit - the error rate for K=6 at this energy is roughly 10^{-14}.

In Figure 4 we have plotted the optimal thresholds, normalized by the signal energy, s, for those error rate curves plotted in Figure 3. As the incident optical energy per bit increases, the normalized optimal threshold increases to unity, which is the curve corresponding to the asymptotically optimal test. Note that the Gaussian MAI, perfect optical-to-electrical approximation predicts a threshold that significantly underestimates the true optimal threshold for those incident optical energies needed to dominate the dark current.

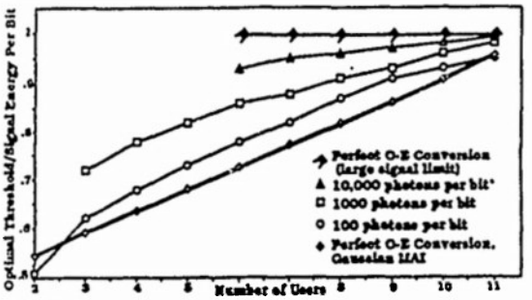

Figure 4. Optimal Thresholds For The Matched-Filter CDMA Receiver

Observe that the asymptotic test yields a more accurate estimate of the optimal threshold for moderate signal energies. Optimal thresholds for large incident optical energies are not plotted for the "error-free" region because they could not be reliably determined due to the vanishing error rate.

References

[1] P.R. Prucnal, M.A. Santoro, T.R. Fan, *"Spread Spectrum Fiber-Optic Local Area Network Using Optical Processing"*, J. Lightwave Tech., Vol. LT-4, No, 5, pp. 547-554, May,1986

[2] J.A. Salehi, C.A. Brackett, *"Fundamental Principles of Fiber Optic Code Division Multiple Access (FO-CDMA)"*, International Conference on Communications, 1987

[3] M.C. Teich, K. Matsuo, B.E.A. Saleh *"Counting Distributions and Error Probabilities for Optical Receivers Incorporating Superlattice Avalanche Photodiodes"*, Electron Devices. ED-33, pp.1475-1488, Oct, 1986

[4] D.L.Snyder, *Random Point Processes,* J. Wiley & Son,1975

[5] A.A. Shaar, P.A.Davies, *"Prime Sequences. Quasi-optimal Sequences For OR Channel Code Division Multiplexing"*, Electronic Letters Vol. 19, No.21,pp. 888-889, 13th, October, 1983

[6] R.F. Serfozo *"Conditional Poisson Processes,"* J. Appl. Prob., 9, pp.288-302 Sept, 1972

[7] J. Grandell, *"Statistical Inference for Double Stochastic Poisson Processes"*, in *Stochastic Point Processes, Ed. P.A.W. Lewis. J. Wiley, New York,* 1972

Optimal Signal Design for Band-Limited PAM Synchronous Multiple-Access Channels

Roger S. Cheng and Sergio Verdú
Department of Electrical Engineering
Princeton University
Princeton, NJ 08544

Abstract

The optimal signal design problem for a band-limited PAM symbol-synchronous Gaussian two-user multiple-access channel is investigated. Using the root-mean-square and the fractional out-of-band energy bandwidth definitions, we find the capacity region of the channel and the signature waveforms to achieve each point inside the capacity region. The optimal pair of signature waveforms are mirror images of each other, and are obtained by minimizing their cross-correlation subject to a fixed finite duration and the bandwidth constraint. The two-user capacity region, in the rms case, is found to contain the capacity region of the two-user strictly band-limited Gaussian channel. This demonstrates the fact that by relaxing constraints in the frequency domain, we can introduce structure (PAM) in the time domain and obtain a larger capacity region.

1. Introduction

The capacity region of the two-user discrete-time Gaussian multiple-access channel

$$y_i = x_{1i} + x_{2i} + n_i$$

where n_i is an i.i.d. Gaussian sequence with variance equal to σ^2 and the energy of each codeword is constrained to satisfy

$$\frac{1}{n}\sum_{i=1}^{n} x_{ki}^2 \le w_k \qquad k = 1, 2$$

is equal to the Cover-Wyner pentagon [1], [2]:

$$C_D = \left\{ (R_1, R_2): \begin{array}{c} 0 \le R_1 \le \frac{1}{2}\log[1 + \frac{w_1}{\sigma^2}] \\ 0 \le R_2 \le \frac{1}{2}\log[1 + \frac{w_2}{\sigma^2}] \\ R_1 + R_2 \le \frac{1}{2}\log[1 + \frac{w_1+w_2}{\sigma^2}] \end{array} \right\} \quad (1)$$

in information units per channel use. Analogously, the capacity region of the continuous-time band-limited channel with noise power spectral density, bandwidth and k^{th} user signal power equal to σ^2, B, and S_k respectively, is given by [2], as (in units per second)

$$C_C = \left\{ (R_1, R_2): \begin{array}{c} 0 \le R_1 \le B\log[1 + \frac{S_1}{2\sigma^2 B}] \\ 0 \le R_2 \le B\log[1 + \frac{S_2}{2\sigma^2 B}] \\ R_1 + R_2 \le B\log[1 + \frac{S_1+S_2}{2\sigma^2 B}] \end{array} \right\} \quad (2)$$

This capacity region is achieved by approximately band-limited and approximately time limited waveforms which have no particular structure. In order to deal with modulation and demodulation schemes with manageable complexity, it is customary in digital communications to introduce structure on the transmitted waveforms by slotting the time domain into intervals of length T and sending a symbol in each slot by means of a digital modulation format such as PAM, PSK, FSK, etc. In the case of PAM (Pulse Amplitude Modulation), the k^{th} user is assigned a fixed deterministic waveform, $s_k(t)$, which is time-limited to $[0, T]$ and is modulated by the information stream. Then, assuming that the transmitters are symbol-synchronous, the PAM two-user multiple-access channel becomes

$$y(t) = \sum_{i=1}^{n} b_1(i)s_1(t - iT) + b_2(i)s_2(t - iT) + n(t) \quad (3)$$

where $n(t)$ is white Gaussian noise with spectral density σ^2 and $\{b_k(i)\}$ is the symbol stream transmitted by the k^{th} user Assuming that, without loss of generality, the signature waveforms have unit energy, the energy constraints on the transmitted waveforms become

$$\frac{1}{n}\sum_{i=1}^{n} b_k^2(i) \le w_k \triangleq TS_k \qquad k = 1, 2 \quad (4)$$

It is easy to show that if $s_1(t) = s_2(t)$, then the capacity of (3) under constraints (4) is equal to the Cover-Wyner pentagon (1) (this result remains true even if the users are completely asynchronous [3].) If the signature waveforms are not necessarily identical, then the Cover-Wyner pentagon generalizes to [4]

$$C_V = \left\{ (R_1, R_2): \begin{array}{c} 0 \le R_1 \le \frac{1}{2}\log[1 + \frac{w_1}{\sigma^2}] \\ 0 \le R_2 \le \frac{1}{2}\log[1 + \frac{w_2}{\sigma^2}] \\ R_1 + R_2 \le \frac{1}{2}\log[1 + \frac{w_1+w_2}{\sigma^2} + \\ \qquad \frac{w_1 w_2}{\sigma^4}(1 - \rho^2)] \end{array} \right\} \quad (5a)$$

in information units per channel use or

$$C_V = \left\{ (R_1, R_2): \begin{array}{c} 0 \le R_1 \le \frac{1}{2T}\log[1 + \frac{TS_1}{\sigma^2}] \\ 0 \le R_2 \le \frac{1}{2T}\log[1 + \frac{TS_2}{\sigma^2}] \\ R_1 + R_2 \le \frac{1}{2T}\log[1 + \frac{TS_1+TS_2}{\sigma^2} + \\ \qquad \frac{T^2 S_1 S_2}{\sigma^4}(1 - \rho^2)] \end{array} \right\} \quad (5b)$$

in information units per second, where $\rho = \int_0^T s_1(t)s_2(t)dt$ is the cross-correlation between the signature waveforms.

A natural question to address is the choice of the unit-energy waveforms $s_1(t)$ and $s_2(t)$ to maximize the capacity region C_V. It is clear that the unconstrained solution is to choose orthogonal signature waveforms. Then, $\rho = 0$, and the multiple-access channel is decoupled into independent single-user channels, and each transmitter can transmit at single-user capacity. However, in practice, there are constraints on the choice of the signals (e.g. in Spread Spectrum CDMA systems, the waveforms may be constrained to be Pseudo Noise shift register sequences of given period,) and it is not always possible to assign orthogonal waveforms for all users. In this paper, we will address the optimization of the signature waveforms and their duration T under bandwidth constraints. Since the signature waveforms are strictly time-limited, they cannot be strictly band-limited, and the need arises to quantify the bandwidth of these signals. There are several established ways to accomplish this [5]. In this paper, we will consider the two bandwidth measures of baseband signals that have received most attention from the information theoretic community: the

This work was partially supported by the office of Naval Research under Contract N00014-87-K-0054

root mean square (rms) bandwidth and the fractional out-of-band energy (fobe) bandwidth.

The rms bandwidth was popularized by Gabor [6] (it is sometimes referred to as Gabor bandwidth) and studied subsequently in [5], [7], and [8]. A finite-energy signal $s(t)$ has rms bandwidth B if its Fourier transform $S(f)$ satisfies

$$\frac{\int_{-\infty}^{\infty} f^2 |S(f)|^2 df}{\int_{-\infty}^{\infty} |S(f)|^2 df} = B^2 \tag{6}$$

i.e. the rms bandwidth is the square root of the "second moment" of the energy spectral density ($|S(f)|^2$) of the normalized signal or, proportional to the square root of the energy of its derivative,

$$\frac{1}{(2\pi)^2} \frac{\int_{-\infty}^{\infty} [\frac{d}{dt} s(t)]^2 dt}{\int_{-\infty}^{\infty} s^2(t) dt} = B^2 \tag{7}$$

The fobe bandwidth has been used in e.g.[5], [8] and is defined as the bandwidth necessary to encompass a given fraction (say α) of the signal energy, i.e. the α-fobe bandwidth is B if

$$\int_{-B}^{B} |S(f)|^2 df = \alpha \int_{-\infty}^{\infty} |S(f)|^2 df \tag{8}$$

Notice that the bandwidth constraints imposed on the signature waveforms will be inherited by the transmitted signals because, as is well known [9], the power spectral density of $\sum_i b_k(i) s_k(t - iT - \tau)$ where τ is uniformly distributed in $[0, T]$ and $\{b_k(i)\}$ is an i.i.d. sequence, is a scaled version of the energy spectral density $|S_k(f)|^2$.

2. Single-user Channel

Before solving for the capacity region of the PAM multiple-access channel under bandwidth constraints, it is enlightening to examine the PAM single-user channel with constrained rms bandwidth. This channel differs from the classical band-limited Gaussian channel in that the allowable transmitted signals 1) have much more structure (PAM) and 2) are rms band-limited but not strictly band-limited. It turns out that the effect of the laxer bandwidth measure cancels the effect of the additional structure imposed on the transmitted signals in the time domain, and the capacity of the channel is given by the celebrated Shannon formula [10].

Theorem 2.1.

The capacity of the single-user PAM white Gaussian channel with noise power spectral density, rms bandwidth and signal power equal to σ^2, B, and S respectively is given by (in units per second)

$$C_S = B \log[1 + \frac{S}{2\sigma^2 B}] \tag{9}$$

Proof.

The single-user PAM white Gaussian channel is a special case of (3):

$$y(t) = \sum_{i=1}^{n} b(i) s(t - iT) + n(t) \tag{10}$$

Assuming that, without loss of generality, $s(t)$ has unit energy, the power constraint becomes

$$\frac{1}{n} \sum_{i=1}^{n} b^2(i) \leq TS \tag{11}$$

and the T-shifts of $s(t)$, $\{s(t - iT)\}_{i=1}^{n}$, form an orthonormal set. The projections of $y(t)$ on this orthonormal set are equal to

$$y(i) = \int_{iT}^{(i+1)T} y(t) s(t - iT) dt \quad i = 1, \ldots, n \tag{12}$$

or, substituting $y(t)$ from (10),

$$y(i) = b(i) + n(i) \tag{13}$$

where $\{n(i)\}$ is an i.i.d. Gaussian sequence with variance equal to σ^2.

The important point to note is that $\{y(i)\}_{i=1}^{n}$ are *sufficient statistics* for the transmitted messages; therefore, the capacity of the PAM channel (10) for a fix T coincides with the capacity of the discrete time memoryless channel (13) with constraint (11), which is given by (e.g. [11]) (in units per second)

$$C_S(T) = \frac{1}{2T} \log[1 + \frac{ST}{\sigma^2}] \tag{14}$$

Since $C_S(T)$ is monotonically decreasing in T, the capacity is maximized by minimizing T. However, due to the rms bandwidth constraint, the value of T cannot be arbitrarily small. Using the fact that the set $\{\sqrt{\frac{2}{T}} \sin(\frac{i\pi t}{T})\}_{i=1}^{\infty}$ is a complete orthonormal set in the space of all rms band-limited signals in $[0, T]$ [7], we can express $s(t)$, as

$$s(t) = \sum_{i=1}^{\infty} d_i \sqrt{\frac{2}{T}} \sin(\frac{i\pi t}{T}) \tag{15}$$

Then, the unit energy assumption and the constraint in the rms bandwidth (7) translate into

$$\sum_{i=1}^{\infty} d_i^2 = 1 \tag{16}$$

and

$$\sum_{i=1}^{\infty} i^2 d_i^2 \leq (2BT)^2 \tag{17}$$

respectively.

The minimum T consistent with (16) and (17) is chosen by taking equality in (17) and minimizing the left hand side of (17) subject to (16). Since

$$1 = \sum_{i=1}^{\infty} d_i^2 \leq \sum_{i=1}^{\infty} i^2 d_i^2 \tag{18}$$

with equality if and only if $d_1 = 1$ and $d_i = 0$ if $1 < i$, it follows that the optimum T is equal to $\frac{1}{2B}$ which upon substitution in (14) results in the desired result. ∎

3. Two-user Channel

We turn our attention to the main results of the paper, namely the optimization of the capacity region of the synchronous PAM channel (5b) with respect to the choice of the signature waveforms, including their duration T. In both the rms and the fobe bandwidth constrained problems, we will solve the problem in two stages:

1. Fix T, and find $\rho^*(TB)$, the minimum absolute cross-correlation, $|\rho|$, achievable under the time-bandwidth constraint (and the optimal waveforms which achieve that ρ.) Then, the capacity region for fixed T is given by C_V in (5b) evaluated at $\rho = \rho^*(TB)$. This is because C_V depends on the signature waveforms only through the rate-sum constraint which is monotonic decreasing in ρ.
2. Take the union of the capacity regions found in the first stage over all T. Note that there is a minimum value of T below which the time-bandwidth product is so small that no waveform can be found to satisfy the bandwidth constraint and therefore, the capacity region is an empty set. Also, there is a maximum value of T above which the allowed time-bandwidth product is so large that orthogonal signals can be assigned to both users, and therefore the capacity region decreases with T beyond that maximum value of T.

Theorem 3.1.

If $TB \geq 0.5$, then the minimum cross-correlation, $\rho_G^*(TB)$, between any two unit-energy signals of duration T and rms bandwidth less than or equal to B is

$$\rho_G^*(TB) = \max\{0, \frac{1}{3}[5 - 8(TB)^2]\}$$

and is achieved by the signature waveforms

$$s_1(t) = \sqrt{\frac{1+\rho_G^*(TB)}{T}} \sin\frac{\pi t}{T} + \sqrt{\frac{1-\rho_G^*(TB)}{T}} \sin\frac{2\pi t}{T}$$

$$s_2(t) = \sqrt{\frac{1+\rho_G^*(TB)}{T}} \sin\frac{\pi t}{T} - \sqrt{\frac{1-\rho_G^*(TB)}{T}} \sin\frac{2\pi t}{T}$$

If $TB < 0.5$, then there exists no signal of duration T and rms bandwidth less than or equal to B.

Proof.

If $TB < 0.5$, we have seen in the proof of Theorem 2.1, that there is no signal of duration T and rms bandwidth less than or equal to B.

If $TB = 0.5$, we have seen that there is only one signal of duration T and rms bandwidth B and is $\sqrt{\frac{2}{T}} \sin\frac{\pi t}{T}$, $t \in [0,T]$. Therefore, the theorem follows immediately when $TB = 0.5$.

If $TB > 0.5$, let $s_1(t)$, $s_2(t)$ he any two unit-energy signals with duration T and rms handwidth B. Using the same complete orthonormal set in the last theorem, we denote the vector $\mathbf{M}(t) = [\sqrt{\frac{2}{T}}\sin(\frac{\pi t}{T}), \sqrt{\frac{2}{T}}\sin(\frac{2\pi t}{T}), \ldots]^T$, $t \in [0,T]$, and express $s_1(t)$ and $s_2(t)$ as

$$s_k(t) = \mathbf{a}_k^T \mathbf{M}(t) \quad k = 1,2 \tag{19}$$

Then, the rms handwidth constraint can be expressed, via (7), as

$$\frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}(\frac{d}{dt}s(t))^2 dt = \frac{1}{(2T)^2}\mathbf{a}_k^T \Pi \mathbf{a}_k \leq B^2 \quad k = 1,2 \tag{20}$$

where $\Pi = \mathrm{diag}[1^2, 2^2, 3^2, \ldots]$. Denoting ρ as the cross-correlation, we can assume that, without loss of generality, $0 \leq \rho$. From the unit energy assumption, we have the cross-correlation matrix, H, as

$$\mathbf{H} = \mathbf{A}\mathbf{A}^T \triangleq \begin{bmatrix}\mathbf{a}_1^T \\ \mathbf{a}_2^T\end{bmatrix}[\mathbf{a}_1 \mathbf{a}_2] = \begin{bmatrix}1 & \rho \\ \rho & 1\end{bmatrix} \tag{21}$$

Since the mapping between $s_k(t)$ and $\mathbf{a}_k$ is an one-to-one mapping, the problem is equivalent to finding the minimum ρ such that there exists A satisfying (20) and (21).

We solve this problem by first giving a lower bound on the cross-correlation and then showing that the lower bound is achievable. Let B_a be the minimum of the sum of the rms bandwidth of M equal energy signals of duration T and correlation matrix, H. B_a is found hy Nuttall [7], as

$$B_a^2 = \frac{1}{(2T)^2}\frac{1}{M}\sum_{l=1}^{r}\mu_l l^2 \tag{22}$$

where each μ_l is the positive eigenvalue of H with $\mu_i \leq \mu_j$ for $j \leq i$, and r is the rank of H.

Appling this result with $M = 2$, $r = 2$ (since $s_1(t) \neq s_2(t)$ implies $\rho \neq 1$) and the correlation matrix H in (21), we get from (20) and (22) that

$$\frac{1}{2(2T)^2}[(1+\rho) + 4(1-\rho)] \leq B^2 \tag{23}$$

where it can be easily verified that $1+\rho$ and $1-\rho$ are eigenvalues of H in (21).

After rearrangement, (23) becomes

$$\frac{1}{3}[5 - 8(TB)^2] \leq \rho \tag{24}$$

Since $s_1(t)$ and $s_2(t)$ are arhitrarily chosen, and ρ helongs to $[0,1]$, we have the lower hound,

$$\max\left\{0, \frac{1}{3}[5 - 8(TB)^2]\right\} \leq \rho_G^*(TB)$$

We now show a signal pair that achieves this lower bound. Stimulated by the fact that the functions $f(t)$ and $f(T-t)$ have the same magnitude spectrum, we consider signature waveforms which are mirror images of each other about $T/2$. Also, we note that $\sin\frac{\pi t}{T}$ is even about $T/2$ while $\sin\frac{2\pi t}{T}$ is odd about $T/2$. Therefore, we assume that the matrix A has the form

$$\mathbf{A} = \begin{bmatrix}\alpha & \sqrt{1-\alpha^2} & 0 & \cdots \\ \alpha & -\sqrt{1-\alpha^2} & 0 & \cdots\end{bmatrix} \tag{25}$$

for some $0 \leq \alpha \leq 1$.

From (20), the rms bandwidth constraint becomes $\sqrt{\frac{4-4(TB)^2}{3}} \leq \alpha$. If we let $\alpha = \sqrt{\frac{4-4(TB)^2}{3}}$ and substitute (25) into (21), we have $\rho = 2\alpha^2 - 1 = \frac{5-8(TB)^2}{3}$. If $\frac{5-8(TB)^2}{3} < 0$, $\sqrt{\frac{4-4(TB)^2}{3}} < \frac{1}{4}$ and we can let $\alpha = \frac{1}{4}$ which gives $\rho = 2\alpha^2 - 1 = 0$. Therefore, we have shown that the lower bound is achievable hy signature waveforms characterized by the matrix A in (25), with $\alpha = \sqrt{\frac{1+\rho^*(TB)}{2}}$. Then, (19) results in the optimal signature waveforms stated in the theorem. ∎

Theorem 3.2.

The capacity region of the two-user PAM white Gaussian multiple-access channel with noise power spectral density, rms bandwidth and signal powers equal to σ^2, B, S_1 and S_2, re-

spectively, is given by

$$C_G = \bigcup_{1\le\gamma\le\sqrt{\frac{5}{2}}} \left\{ (R_1,R_2): \begin{array}{c} 0 \le R_1 \le \frac{B}{\gamma}\log[1+\frac{S_1\gamma}{2\sigma^2 B}] \\ 0 \le R_2 \le \frac{B}{\gamma}\log[1+\frac{S_2\gamma}{2\sigma^2 B}] \\ R_1+R_2 \le \frac{B}{\gamma}\log[1+\frac{(S_1+S_2)\gamma}{2\sigma^2 B}+ \\ \frac{S_1S_2\gamma^2}{4\sigma^4B^2}(1-\frac{4}{9}(\frac{5}{2}-\gamma^2)^2)] \end{array} \right\} \tag{26}$$

Proof.

Recall that the capacity region, C_G, is the union of C_V in (5b) evaluated at $\rho^*(TB)$ over T. We proceed to find the range of T of interest. From the last theorem, if $TB < 0.5$, no signature waveforms can be found to satisfy the constraints and the capacity region is an empty set. Also, if $TB \ge \sqrt{\frac{5}{8}}$, $\rho^*(TB) = 0$, and the capacity region for fixed T is a pentagon which is monotonic decreasing in T. Therefore, the range of T in interest is the interval $[\frac{1}{2B}, \frac{1}{2B}\sqrt{\frac{5}{2}}]$. Denoting $2TB$ by γ, and substituting γ into C_V in (5b), we have, after taking the union, C_G in the theorem. ∎

At a first glance, it seems that there is a conflict with Theorem 2.1 since the total capacity of C_G is larger than the single-user capacity of an rms band-limited channel with power constraint $S_1 + S_2$. However, the signal transmitted over the channel in the two-user case is a sum of two PAM signals and, in general, it is no longer a PAM signal since the signals in different time slots need not have the same shape.

Figure 1 shows the capacity region of the rms band-limited PAM two-user channel, C_G and the strictly band-limited two-user channel, C_C. In contrast to the single-user case where they coincide, C_G is a subset of C_C. It can also be seen from (26) and (2) that C_G is the pentagon inside the union in (26) when $\gamma = 1$. However, by increasing γ, we trade off the decrease in the single-user rate by the increase in the rate sum, such that the union gives a larger capacity region, C_G. This indicates that, in the two-user case, the laxer bandwidth constraint more than offsets the additional structure (PAM) in the time domain.

Figure 2 and 3 show the signature waveforms which achieve the boundary points of the capacity region for two different time-bandwidth products. The signature waveforms are mirror images of each other and as γ increases, they become more asymmetric so as to decrease the cross-correlation while maintaining the same rms bandwidth.

Finally, although the union in Theorem 3.2 is taken over γ in the interval $[1, \sqrt{\frac{5}{2}}]$, not every γ in that interval achieves some boundary points of C_G. The set of values of γ that achieves boundary points of C_G is a function of the signal-to-noise ratios, $\frac{S_k}{2\sigma^2 B}$, $k = 1,2$. According to Figure 1, the boundary points in the segments AB and EF are achieved by $\gamma = 1$, while those in the segment CD are achieved by some $\gamma_{\max}$ in $[1, \sqrt{\frac{5}{2}}]$ depending on the signal-to-noise ratios. The boundary points in BC and DE are achieved by $1 \le \gamma \le \gamma_{\max}$.

We now proceed to the optimal signal design problem under α-fobe bandwidth constraint. Denote the prolate spheroidal wave functions ([12], [13], and [14]) as $\psi_i(TB,t)$ and the associated eigenvalues as $\lambda_i(TB)$, i.e.

$$\lambda_i(TB)\psi_i(TB,t) = \int_{-\frac{1}{2}}^{\frac{1}{2}} \psi_i(TB,\tau)\frac{\sin[2\pi TB(t-\tau)]}{\pi(t-\tau)}d\tau$$

for $i = 0,1,2,\ldots$ and $\lambda_0(TB) > \lambda_1(TB) > \lambda_2(TB) > \cdots$. It is known [8] that $\psi_0(TB, \frac{t}{T} - \frac{1}{2})$ and $\psi_1(TB, \frac{t}{T} - \frac{1}{2})$ are even and odd about $\frac{T}{2}$ respectively and the set $\left\{\frac{1}{\sqrt{\lambda_i(TB)}}\psi_i(TB, \frac{t}{T} - \frac{1}{2})\right\}$ forms a complete orthonormal set in $[0,T]$. Also, $\lambda_0(TB)$ and $\lambda_0(TB) + \lambda_1(TB)$ are continuous and monotonic increasing in TB (Figure 4).

Theorem 3.3.

For any $0 < \alpha < 1$,

If $TB \ge \lambda_0^{-1}(\alpha)$, then the minimum cross-correlation, $\rho_F^*(TB)$, between any two unit-energy signals of duration T, and α-fobe bandwidth less than or equal to B is

$$\rho_F^*(TB) = \max\left\{0, \frac{2\alpha - \lambda_0(TB) - \lambda_1(TB)}{\lambda_0(TB) - \lambda_1(TB)}\right\}$$

and is achieved by the signature waveforms

$$s_1(t) = \sqrt{\frac{1+\rho_F^*(TB)}{2\lambda_0(TB)}}\psi_0(TB, \tfrac{t}{T} - \tfrac{1}{2}) + \sqrt{\frac{1-\rho_F^*(TB)}{2\lambda_1(TB)}}\psi_1(TB, \tfrac{t}{T} - \tfrac{1}{2})$$

$$s_2(t) = \sqrt{\frac{1+\rho_F^*(TB)}{2\lambda_0(TB)}}\psi_0(TB, \tfrac{t}{T} - \tfrac{1}{2}) - \sqrt{\frac{1-\rho_F^*(TB)}{2\lambda_1(TB)}}\psi_1(TB, \tfrac{t}{T} - \tfrac{1}{2})$$

If $TB < \lambda_0^{-1}(\alpha)$, then there exists no signal of duration T and α-fobe bandwidth less than or equal to B.

Proof.

As in Theorem 3.1, we would like to find a suitable complete orthonormal set in $[0,T]$. To that end, we rewrite the definition of α-fobe bandwidth as

$$\begin{aligned} \alpha &\le \int_{-B}^{B} |S(f)|^2 df \\ &= \int_0^T \int_0^T s(t)s(\tau)\int_{-B}^{B} e^{j2\pi f(t-\tau)} df\, dt\, d\tau \\ &= \int_0^1 \int_0^1 s(\frac{t}{T})\frac{\sin(2\pi TB(t-\tau))}{\pi(t-\tau)} s(\frac{\tau}{T})\, dt\, d\tau \end{aligned} \tag{27}$$

Since the prolate spheroidal wave functions are eigenfunctions of the kernel $2TB\text{sinc}(t-\tau)$, a good choice for the complete orthonormal set will be the set of all prolate spheroidal wave functions.

For notational convenience, we will drop the explicit dependence on TB of the eigenvalues of the prolate spheroidal wave functions. If $TB \ge \lambda_0^{-1}(\alpha)$, we can express any $s_1(t)$ and $s_2(t)$ in terms of $\Psi(t) = [\frac{1}{\sqrt{\lambda_0}}\psi_0(TB, \frac{t}{T} - \frac{1}{2}), \frac{1}{\sqrt{\lambda_1}}\psi_1(TB, \frac{t}{T} - \frac{1}{2}), \ldots]$, $t \in [0,T]$, as

$$s_k(t) = \mathbf{a}_k^T\Psi(t) \quad k = 1,2 \tag{28}$$

Using (27) and (28), we have

$$\alpha \le \int_{-B}^{B} |S(f)|^2 df = \mathbf{a}_k^T\Lambda\mathbf{a}_k = \text{tr}(\Lambda\mathbf{a}_k\mathbf{a}_k^T) \qquad k = 1,2 \tag{29}$$

where $\Lambda = \text{diag}[\lambda_0, \lambda_1, \lambda_2, \ldots]$. Also, the cross-correlation matrix, $\mathbf{H}$, is

$$\mathbf{H} = \mathbf{A}\mathbf{A}^T = \begin{bmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \end{bmatrix}[\mathbf{a}_1\mathbf{a}_2] = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \tag{30}$$

Similar to the rms case, we find the lower bound by maximizing the average over $k = 1, 2$ of the right hand side of (29). Rewriting the average, we have

$$\frac{1}{2}\sum_{k=1}^{2}\mathbf{a}_k^T\Lambda^T\mathbf{a}_k = \frac{1}{2}\mathrm{tr}(\Lambda\mathbf{A}^T\mathbf{A}) = \frac{1}{2}\mathrm{tr}(\Lambda\mathbf{P}_\mathbf{A}\Xi_\mathbf{A}\mathbf{P}_\mathbf{A}^T) = \frac{1}{2}\mathrm{tr}(\Xi_\mathbf{A}\mathbf{P}_\mathbf{A}^T\Lambda\mathbf{P}_\mathbf{A}) \tag{31}$$

where $\mathbf{A}^T\mathbf{A}$ is diagonalized by the orthonormal matrix, $\mathbf{P}_\mathbf{A}$, and $\Xi_\Lambda = \mathrm{diag}[\xi_1 \quad \xi_2 \quad 0 \quad \ldots]$. Since the eigenvalues of $\mathbf{A}\mathbf{A}^T$ and $\mathbf{A}^T\mathbf{A}$ are the same, we have $\xi_1 = 1+\rho$ and $\xi_2 = 1-\rho$.

Now, let's denote $\mathbf{P}$ as the 2x∞ matrix formed by taking only the first two rows of $\mathbf{P}_\mathbf{A}^T$, and Ξ as $\mathrm{diag}[\xi_1 \quad \xi_2]$. Then, the maximum of the average is

$$\max_{\mathbf{P}\mathbf{P}^T=\mathbf{I}_{2\times 2}} \mathrm{tr}(\Xi\mathbf{P}\Lambda\mathbf{P}^T) \tag{32}$$

We will solve the maximization problem using the Lagrange multiplier method. We form the Lagrangian,

$$\sum_{k=1}^{2}\xi_k\mathbf{p}_k^T\Lambda\mathbf{p}_k + \sum_{k=1}^{2}\sum_{n=1}^{k} z_{nk}(\mathbf{p}_k^T\mathbf{p}_n - \delta_{nk}) \tag{33}$$

where $\mathbf{p}_k^T$ is the k^{th} row of $\mathbf{P}$. Taking derivative with respect to $\mathbf{p}_k$, we have

$$\xi_1\Lambda\mathbf{p}_1 + z_{11}\mathbf{p}_1 + 2z_{12}\mathbf{p}_2 = 0 \tag{34}$$

and

$$\xi_2\Lambda\mathbf{p}_2 + z_{22}\mathbf{p}_2 + 2z_{12}\mathbf{p}_1 = 0 \tag{35}$$

If we pre-multiply (34) by $\mathbf{p}_2^T$ and (35) by $\mathbf{p}_1^T$, we have $z_{12} = 0$ since $\xi_1 \neq \xi_2$; therefore, from (34) and (35), $\mathbf{p}_1$ and $\mathbf{p}_2$ are eigenvectors of Λ. Since Λ is diagonal and the diagonal elements are distinct and decreasing down the diagonal axis, we have

$$\mathbf{P} = [\mathbf{I}_{2\mathbf{x}2} \quad 0] \tag{36}$$

Substituting back into (31), we show that the maximum value of (31) is $\lambda_0\frac{1+\rho}{2} + \lambda_1\frac{1-\rho}{2}$. Comparing to (29), we have

$$\alpha \leq \frac{1}{2}[(\lambda_0+\lambda_1) + \rho(\lambda_0-\lambda_1)]$$

or, together with $0 \leq \rho \leq 1$,

$$\rho \geq \max\left\{0, \frac{2\alpha-\lambda_0-\lambda_1}{\lambda_0-\lambda_1}\right\}$$

The achievability of the lower bound can be verified, as in the rms case, by letting

$$\mathbf{A} = \begin{bmatrix} \sqrt{\frac{1+\rho}{2}} & \sqrt{\frac{1-\rho}{2}} & 0 & \cdots \\ \sqrt{\frac{1+\rho}{2}} & -\sqrt{\frac{1-\rho}{2}} & 0 & \cdots \end{bmatrix} \tag{37}$$

which corresponds to the optimal set of signals stated in the theorem.

The proof of the second part of the theorem ($TB < \lambda_0^{-1}(\alpha)$) can be found in [13, p.54]. ∎

Theorem 3.4.

The capacity region of the two-user PAM white Gaussian multiple-access channel with noise power spectral density, α-fobe bandwidth and signal powers equal to σ^2, B, S_1, and S_2, respectively, is given by

$$C_F = \bigcup_{\gamma_{\min}\leq\gamma\leq\gamma_{\max}} \left\{ (R_1,R_2): \begin{array}{l} 0 \leq R_1 \leq \frac{B}{\gamma}\log[1+\frac{S_1\gamma}{2\sigma^2 B}] \\ 0 \leq R_2 \leq \frac{B}{\gamma}\log[1+\frac{S_2\gamma}{2\sigma^2 B}] \\ R_1+R_2 \leq \frac{B}{\gamma}\log[1+\frac{(S_1+S_2)\gamma}{2\sigma^2 B} + \\ \quad \frac{S_1S_2\gamma^2}{4\sigma^4B^2}[1-(\frac{2\alpha-\lambda_0(\frac{\gamma}{2})-\lambda_1(\frac{\gamma}{2})}{\lambda_0(\frac{\gamma}{2})-\lambda_1(\frac{\gamma}{2})})^2]] \end{array} \right\}$$

where $\alpha = \lambda_0(\frac{\gamma_{\min}}{2}) = \frac{1}{2}[\lambda_0(\frac{\gamma_{\max}}{2}) + \lambda_1(\frac{\gamma_{\max}}{2})]$.

Proof.

The proof is very similar to that in Theorem 3.2 where $\gamma = 2TB$. The lower limit of γ is carried over from Theorem 3.3, while $\gamma_{\max}$ is the smallest γ such that $\rho_F^*(\frac{\gamma}{2}) = 0$. ∎

Notice that the range of γ in taking the union is only a function of α. In Figure 4, we show $\lambda_0(TB)$ and $\frac{1}{2}[\lambda_0(TB)+\lambda_1(TB)]$ vs the time-bandwidth product, and $\gamma_{\min}$ and $\gamma_{\max}$ can be obtained directly from the figure. Also, Figure 5 shows the capacity region, C_F, with the capacity region of the strictly band-limited channel, C_C. Similar comments to those we made in the rms case apply to the values of γ that achieve the boundary points in the capacity region. However, we see that for sufficiently high α, C_F does not contain C_C in contrast to the rms case.

Finally, in Figure 6, we show the signature waveforms which are, as expected, mirror images of each other. However, in contrast to the rms case where the signature waveforms must be zero at the end points to have finite rms bandwidth, the transmitted signal waveform in the α-fobe case may have jumps at $t = iT$.

References

[1] Cover, T., "Some advances in broadcast channels," *Advances in Communication Systems*, vol. 4, pp. 229-260, Academic Press, New York, 1975.

[2] Wyner, A. D., "Recent results in the Shannon theory," *IEEE Transactions on Information Theory*, vol. IT-20, January, 1974.

[3] Verdú, S., "The capacity region of the symbol-asynchronous Gaussian multiple-access channel," *IEEE Transactions on Information Theory*, to appear.

[4] Verdú, S., "Capacity region of Gaussian CDMA channels: the symbol-synchronous case," *Proc. of the Twenty-fourth Allerton Conference on Communication, Control and Computing*, Allerton, IL, pp. 1025-1034, October, 1986.

[5] Amoroso, F., "The bandwidth of digital data signals," *IEEE Communication Magazine*, vol. 18, pp. 13-24, November 1980.

[6] Gabor, D., "Theory of communication," *J. IEE*, vol. 93, pt. 3, pp. 429-457, November 1946.

[7] Nuttall, A. H., "Minimum rms bandwidth of M time-limited signals with specified code or correlation matrix,"

IEEE Transactions on Information Theory, vol. IT-14, pp. 699-707, September 1968.

[8] Narayan, P. and D. Snyder, "Signal set design for band-limited memoryless multiple-access channels with soft decision demodulation," *IEEE Transactions on Information Theory*, vol. IT-33, pp. 539-556, July, 1987.

[9] Lee, E. A. and D. G. Messerschmitt, *Digital Communication*, Kluwer Academic Publishers, Boston, 1988.

[10] Shannon, C. E., "A mathematical theory of communication II," *Bell System Technical Journal*, vol. 27, pp. 623-656, October 1948.

[11] Gallager, R. G., *Information theory and reliable communication*, New York: John Wiley & Son, 1968.

[12] Flammer C., *Spheroidal wave functions*, Stanford University Press, Stanford, California, 1957.

[13] Slepian, D. and H. O. Pollak, "Prolate spheroidal wave functions, fourier analysis and uncertainty I," *Bell System Technical Journal*, vol. 40, pp. 43-63, January, 1961.

[14] Landau, H. J. and H. O. Pollak, "Prolate spheroidal wave functions, fourier analysis and uncertainty II," *Bell System Technical Journal*, vol. 40, pp. 65-84, January, 1961.

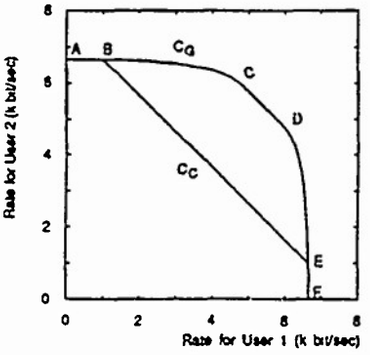

Figure 1. Capacity Regions in the rms case for SNR1=SNR2=20db, B=1khz

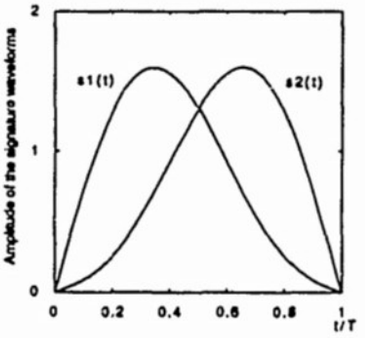

Figure 2. Signature waveforms for two-user rms band-limited channel, TB=0.6

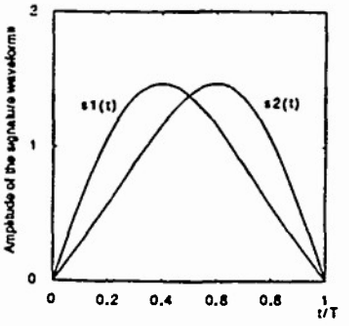

Figure 3. Signature waveforms for two-user rms band-limited channel, TB=0.525

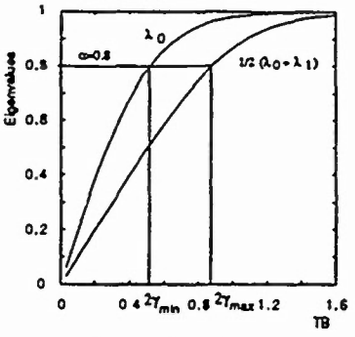

Figure 4. Eigenvalues vs. time-bandwidth product

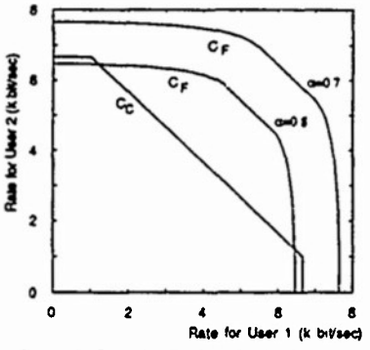

Figure 5. Capacity Regions in the fobe case for SNR1=SNR2=20db, B=1khz

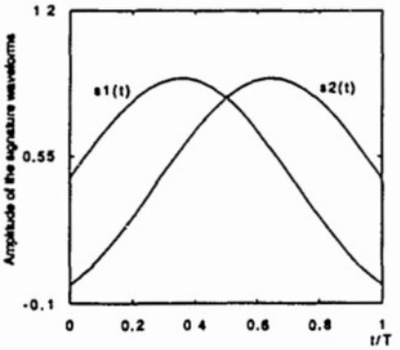

Figure 6. Signature waveforms for two-user 0.8-fobe band-limited channel, TB=0.955

"OTAL CAPACITY OF THE RMS BANDLIMITED K-USER PAM SYN- :HRONOUS CHANNEL

ιOGER S. CHENG & SERGIO VERDÚ †
)epartment of Electrical Engineering
'rinceton University, Princeton, NJ 08544

ABSTRACT

'ontinuous-time additive white Gaussian noise channels with strictly time-limited and root mean square (RMS) bandlimited inputs are studied. RMS bandwidth is equal to the normalized second moment of the spectrum, which has proved to be a useful and analytically tractable measure of the bandwidth of strictly time-limited waveforms.

We 'nd the Total Capacity (TC) of the K-user channel under total power and power-weig..t ..d average RMS bandwidth constraints. A lower bound to the TC under equal-power constraint is obtained. Total Capacity Ratio (TCR) is defined as the ratio of the K-user TC to K times the single-user capacity. Power (Bandwidth) efficiency is defined as the ratio of the effective power (bandwidth) to the actual power (bandwidth). The effective power (bandwidth) is the corresponding power (bandwidth) needed for a single user channel to achieve the same capacity. We find lower bounds to the TCR and efficiencies which indicate that savings in bandwidth compared to the FDMA scheme can be achieved by the CDMA scheme at the expense of more complicated decoding hardware.

1. INTRODUCTION

In this paper, we deal with the continuous-time Pulse Amplitude Modulation (PAM) Gaussian multiple-access channel (MAC). Each user is assigned a fixed deterministic continuous-time signature waveform, $s_k(t)$, which is time-limited to $[0,T]$ and is modulated linearly by the information stream. Assuming that the transmitters are symbol-synchronous, the channel can be expressed as

$$y(t) = \sum_{i=1}^{n} \sum_{k=1}^{K} b_k(i) s_k(t - iT) + n(t) \tag{1}$$

where $n(t)$ is white Gaussian noise with spectral density, $\frac{N_0}{2}$ and $\{b_k(i)\}$ is the symbol stream transmitted by the k^{th} user.

The capacity region of this channel has been found by Verdú [1] [2]. Denoting $\mathbf{W}$ and $\mathbf{H}$ as the diagonal matrix with the users' powers as its diagonal entries, and the cross-correlation matrix of the normalized signature waveforms, respectively, the capacity region is expressed as

$$\mathrm{C}_V = \left\{ (R_1, R_2, \ldots, R_K) : \sum_{j \in J} R_j \leq \frac{1}{2T} \log[\det(\mathrm{I}_{|J|} + \frac{2T}{N_0} \mathbf{W}_J \mathbf{H}_J)] \quad \forall J \subset \{1, \ldots, K\} \right\} \tag{2}$$

This work was supported by the office of Naval Research under Contract N00014-87-K-0054 ..nd by the National Science Foundation under PYI Grant ECSE-8857689

where $\mathbf{A}_J$ is the $|J|\times|J|$ matrix formed by the j^{th} row and column of $\mathbf{A}$ for all $j \in J$. It is clear that, without other constraints, the capacity region is maximized by orthogonal signature waveforms. However, under bandwidth constraints, orthogonal signature waveforms are not necessarily optimal since orthogonality can only be achieved by lowering the symbol rate, $1/T$.

There are many different bandwidth definitions [3]. In this paper, we concentrate on the root mean square (RMS) bandwidth because it is analytically tractable and can be applied to strictly time-limited signals. The RMS bandwidth was introduced by Gabor [4] and studied subsequently in [3], [5] and [6]. It is the square root of the second moment of the energy spectral density ($|S_k(f)|^2$) of the normalized signal which is proportional to the square root of the energy of its derivative.

In the two-user case, the capacity region of the RMS bandlimited PAM channel has been found in [7] and the total capacity (the maximum rate sum over the capacity region) is larger than the single-user capacity with the power equal to the sum of the users' powers. The gain in the total capacity from the single-user to the two-user case can be explained by the increase in the dimensionality of the signal set. We can consider the transmitted signal in a symbol interval as a signal drawn from a signal set. Then, the signal set in the single-user and the two-user case are one-dimensional and two-dimensional, respectively. From this viewpoint, it is easy to see that the total capacity increases as the number of users increases while the total power remains constant.

In this paper, we find the total capacity (TC) of the K-user channel under the total power constraint

$$\mathrm{tr}(\mathbf{W}) \le W \tag{3}$$

and the power-weighted averaged RMS bandwidth constraint

$$\frac{1}{\mathrm{tr}(\mathbf{W})}\sum_{k=1}^{K}\mathbf{W}_{kk}\int_{-\infty}^{\infty} f^2\,|S_k(f)|^2\,df \le B^2 \tag{4}$$

The power constraint is placed on the total power instead of the individual power since the later requires finding all possible sets of eigenvalues of a positive definite matrix with fixed diagonal entries which is, in general, intractable. The bandwidth constraint is justified because the power-weighted average RMS bandwidth is the RMS bandwidth of the power spectral density of the transmitted signal.

Several performance measures, Total Capacity Ratio, Power efficiency and Bandwidth efficiency, are defined and analyzed. Bounds and limiting values of these measures are also obtained.

2. TOTAL CAPACITY

Theorem 2.1.

The Total Capacity of the K-user RMS bandlimited PAM Gaussian MAC with total power and power-weighted average RMS bandwidth constraints is

$$\mathrm{C}(B,K,\Lambda) = \max\left\{\frac{B}{\gamma}\sum_{n=1}^{N}\log[1+h_n(\lambda)]\right\}\log e \tag{5}$$

where the maximization is over $1 \le N \le K$, $\frac{\gamma K \Lambda}{N} \le \lambda$, *and* $1 \le \gamma \le \sqrt{f_N}$ *such that* $\gamma = \sqrt{f_N}$ *iff* $\lambda = \frac{\gamma K \Lambda}{N}$, *and*

$$\sum_{n=1}^{N} h_n(\lambda) = \gamma K \Lambda \tag{6}$$

where

$$h_n(\lambda) \triangleq \frac{N\lambda f_N + \gamma K \Lambda[\lambda(\gamma^2 - 1) - 1] - n^2(N\lambda - \gamma K \Lambda)}{N(f_N - 1 - \lambda) + \gamma^3 K \Lambda + n^2(N\lambda - \gamma K \Lambda)} > 0 \tag{7}$$

$$f_N \triangleq \frac{1}{N} \sum_{n=1}^{N} n^2 = \frac{1}{6}(N+1)(2N+1) \tag{8}$$

and the average signal-to-noise ratio is denoted by

$$\Lambda \triangleq \frac{W}{K N_0 B} \tag{9}$$

Proof.

Since the signature waveforms are RMS bandlimited, and the set $\{\phi_i(t,T)\}_{i=1}^{\infty}$ where

$$\phi_i(t,T) \triangleq \begin{cases} \sqrt{\frac{2}{T}} \sin(\frac{\pi i t}{T}) & \text{if } t \in [0,T]; \\ 0 & \text{otherwise.} \end{cases} \tag{10}$$

forms a complete orthonormal basis for all RMS bandlimited signals, we can write

$$\begin{bmatrix} s_1(t) \\ s_2(t) \\ \vdots \\ s_K(t) \end{bmatrix} = \mathbf{A}^T \Phi(t,T) \tag{11}$$

where $\Phi(t,T) \triangleq [\phi_1(t,T)\ \phi_2(t,T)\ \ldots]^T$. Then, the power-weighted average RMS bandwidth constraint can be written, via the Parseval's theorem, as

$$\begin{aligned} \frac{1}{\mathrm{tr}(\mathbf{W})} \sum_{k=1}^{K} \mathbf{W}_{kk} \int_0^T [\frac{d}{dt} s_k(t)]^2 \, dt &= \frac{1}{(2T)^2 \mathrm{tr}(\mathbf{W})} \mathrm{tr}(\mathbf{W}\mathbf{A}^T \mathbf{\Pi} \mathbf{A}) \\ &= \frac{1}{(2T)^2 \mathrm{tr}(\mathbf{W})} \mathrm{tr}(\mathbf{\Pi} \mathbf{A} \mathbf{W} \mathbf{A}^T) \\ &\le B^2 \end{aligned} \tag{12}$$

From the capacity region in (2), it is clear that the total capacity is maximized when $\mathrm{tr}(\mathbf{W}) = W$. We denote the time-bandwidth product by $\gamma \triangleq 2BT$, the average signal-to-noise ratio by $\Lambda \triangleq \frac{W}{K N_0 B}$, and the eigenvalues of $\frac{2T}{N_0}\mathbf{W}\mathbf{H}$ by λ_k such that $\lambda_j \le \lambda_i, \forall i < j \le K$. Then, the total capacity becomes

$$\mathrm{TC_V} = \frac{B}{\gamma} \sum_{k=1}^{K} \log[1 + \lambda_k] \tag{13}$$

and the power constraint becomes

$$\sum_{k=1}^{K} \lambda_k = \gamma K \Lambda \tag{14}$$

Since the eigenvalues of $\frac{2T}{N_0}\mathbf{WH}$ are also the eigenvalues of $\frac{2T}{N_0}\mathbf{AWA}^T$, and once $\{\lambda_k\}_{k=1}^K$ are fixed, the left hand side of (12) is minimized when $\mathbf{AWA}^T$ is diagonal with decreasing diagonal entries, we can rewrite (12) as

$$\sum_{k=1}^{K} k^2\lambda_k \leq \gamma^3 K\Lambda \tag{15}$$

For fixed T, the total capacity is found by maximizing (13) over all $\lambda_k \geq 0$, $k = 1,\ldots,K$ under the constraints (14) and (15). Using the Kuhn-Tucker Theorem, we form the Lagrangian

$$-\sum_{k=1}^{K}\log[1+\lambda_k] + x(\sum_{k=1}^{K}\lambda_k - \gamma K\Lambda) + y(\sum_{k=1}^{K}k^2\lambda_k - \gamma^3 K\Lambda) \tag{16}$$

and obtain the necessary conditions:

$$\lambda_n \;=\; \frac{1}{x+yn^2} - 1 \;>\; 0 \qquad n = 1,\ldots,N. \tag{17}$$

and $\lambda_n = 0$ for all $n > N$,

$$y(\sum_{n=1}^{N} n^2\lambda_n - \gamma^3 K\Lambda) = 0 \tag{18}$$

and $0 \leq y$.

Rewriting (17) as $(x+yn^2)(1+\lambda_n) = 1$, and summing over all n, we have, from (14) and (18),

$$(N+\gamma K\Lambda)x + (Nf_N + \gamma^3 K\Lambda)y = N \tag{19}$$

Particularizing (17) to $n = 1$, and substituting in (19), we have

$$y = \frac{N\lambda_1 - \gamma K\Lambda}{(1+\lambda_1)(N(f_N-1)+\gamma K\Lambda(\gamma^2-1))} \tag{20}$$

and

$$x = \frac{N(f_N - 1 - \lambda_1) + \gamma^3 K\Lambda}{(1+\lambda_1)(N(f_N-1)+\gamma K\Lambda(\gamma^2-1))} \tag{21}$$

Substituting (20) and (21) into (17), and denoting λ_1 by λ and λ_n by $h_n(\lambda)$, we have (7), and the power constraint in (14) becomes (6).

When $y = 0$, $\lambda = h_n(\lambda) = \frac{\gamma K\Lambda}{N}$ for all $n = 1,\ldots,N$. Upon substituting into (15), we have $\sqrt{f_N} \leq \gamma$. Since the total capacity becomes $\frac{BN}{\gamma}\log[1+\frac{\gamma K\Lambda}{N}]$ which is monotonically decreasing in γ, the optimal γ is equal to $\sqrt{f_N}$ and (15) is satisfied with equality. If we rewrite (7) and sum up over all n, we have

$$(N\lambda_1 - \gamma K\Lambda)(\sum_{n=1}^{N} n^2 h_n(\lambda) - \gamma^3 K\Lambda) = 0 \tag{22}$$

When $0 < y$, $\frac{\gamma K\Lambda}{N} < \lambda$, and from (22), (15) is again satisfied with equality. Therefore, if we require $\gamma = \sqrt{f_N}$ iff $\lambda = \frac{\gamma K\Lambda}{N}$, (15) is superfluous. Finally, specifying the range of γ and λ and the condition that $\gamma = \sqrt{f_N}$ iff $\lambda = \frac{\gamma K\Lambda}{N}$, we have the desired result. ∎

This theorem gives the exact calculation needed for the TC. The main reason why we cannot obtain a simpler solution is the lack of a closed form expression of $\sum_{n=1}^{N}\frac{1}{a+bn^2}$. However,

despite the complicated expression, the TC can be computed once the average signal-to-noise ratio, Λ, and K is given. In Figure 1, we show the TC with different values of K and Λ.

For a given $\mathbf{W}$, we show that any set of signature waveforms, with $\mathbf{A}$ such that $\mathbf{AWA}^T$ is a diagonal matrix with the n^{th} diagonal entry equal to $h_n(\lambda)$, is optimal. However, such an $\mathbf{A}$ does not always exist for any arbitarily given $\mathbf{W}$. For fixed total power, W, finding the set of $\mathbf{W}$ where $\mathbf{A}$ exists is equivalent to finding the possible set of diagonal entries of a positive definite matrix with fixed eigenvalues, which seems intractable. Reversing the problem, one may want to fix the $\mathbf{W}$ and find the total capacity. In general, this is equivalent to finding the possible set of eigenvalues of a positive definite matrix with fixed diagonal entries, which is again intractable.

In the following theorem, we give a lower bound to the TC in the equal-power case where $\mathbf{W} = \frac{W}{K}\mathbf{I}$. Clearly, this is also a lower bound to the capacity of the original channel with the total power constraint in Theorem 2.1.

Theorem 2.2.

The lower bound to the Total Capacity when the users' powers are the same is

$$\mathrm{TC}_{\mathrm{EP}}(B,K,\Lambda) \geq \max_{1\leq\gamma\leq\sqrt{f_K}} \frac{B}{\gamma}\log\left\{[1+\gamma\Lambda\frac{(\gamma^2-1)+K(f_K-\gamma^2)}{f_K-1}][1+\gamma\Lambda\frac{\gamma^2-1}{f_K-1}]^{K-1}\right\} \tag{23}$$

Proof.

The lower bound is found by exhibiting a symmetric positive definite matrix $\mathbf{H}$, such that the total capacity for that particular signature waveform set is easy to find. We let $\mathbf{H}$ be

$$\mathbf{H} = \begin{bmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \rho \\ \rho & \cdots & \rho & 1 \end{bmatrix} \tag{24}$$

the eigenvalues of $\mathbf{H}$ with $0 \leq \rho \leq 1$ to be be specified in the sequel are $1+(K-1)\rho$ and $1-\rho$ with multiplicity $K-1$. Then, the total capacity under the equal-power constraint becomes

$$\begin{aligned} \mathrm{TC}_{\mathrm{V}} &= \frac{B}{\gamma}\log[\det(\mathbf{I}_K + \gamma\Lambda\mathbf{H})] \\ &= \frac{B}{\gamma}\log\left\{[1+\gamma\Lambda(1+(K-1)\rho)][1+\gamma\Lambda(1-\rho)]^{K-1}\right\} \end{aligned} \tag{25}$$

while the bandwidth constraint (15) becomes

$$\rho \geq \frac{f_K-\gamma^2}{f_K-1} \tag{26}$$

Since (25) is monotonically decreasing in ρ when $0 \leq \rho \leq 1$, the TC is maximized when ρ achieves equality in (26). Substituting ρ from (26) with equality into (25), and maximizing over all γ, we have (23). ∎

In Figure 1, we plot the lower bound to the TC_{EP} for different values of K and Λ. Since the TC under the total power constraint serves as an upper bound to the TC_{EP}, Figure 1 gives a tight upper and lower bound to the Total Capacity of a equal-power constrained channel, for moderate number of users.

As a performance measure, we define the Total Capacity Ratio (TCR) as the ratio of the K-user TC to K times the single-user capacity with the same RMS bandwidth and average signal-to-noise ratio constraints. Since the single-user capacity of a RMS bandlimited PAM channel is equal to $B\log[1+\Lambda]$ (see [7]), the TCR can be written as

$$\mathrm{TCR}(K,\Lambda) \triangleq \frac{\mathrm{TC}(B,K,\Lambda)}{KB\log[1+\Lambda]} \tag{27}$$

The TCR gives the ratio of the capacity available to an average user (when the channel is shared by K users) to the single-user capacity. In other words, it measures, from the user's viewpoint, the ratio of the average user capacity in a multi-user channel to the capacity in a single-user channel. Notice that the TCR depends only on K and Λ, and is independent of B. Using the lower bound in Theorem 2.2, we obtain a lower bound to the TCR under the equal-power constraint for all signal-to-noise ratios.

Corollary 2.1.

A lower bound to the TCR under the equal-power constraint for all signal-to-noise ratio is

$$\mathrm{TCR}(K,\Lambda) \geq \frac{1}{\gamma} \tag{28}$$

where γ is the positive real root of the equation

$$\gamma(\gamma^2-1) = f_K - 1 \tag{29}$$

Proof.

In order to obtain (28), we simply substitute γ from (29) into (23) and (27). Since there is one and only one real positive solution in (29), there is no ambiguity in the value of γ. ∎

In Figure 2, we show the TCR under the total power constraint and the lower bound to the TCR under the equal-power constraint for different number of users and different average signal-to-noise ratios.

3. EFFICIENCIES

The TCR gives the performance degradation, from the user's viewpoint, when a bandlimited channel is shared by K users instead of a single user. A natural question to be asked is "How to maintain the same rate in the presence of other users?" If we want to maintain the same information rate, we have to modify some of the parameters. In the following, we will analyze two alternatives. First, we increase the signal-to-noise ratio by increasing the power while the bandwidth remains constant. Second, we increase the bandwidth of the channel while the power of each user remains the same.

The power efficiency, denoted by $\eta_P(K,\Lambda)$, is defined as

$$\eta_P(K,\Lambda) \triangleq \frac{e^{\frac{TC(B,K,\Lambda)}{BK}} - 1}{\Lambda} \tag{30}$$

or, equivalently, implicitly as

$$TC(B,K,\Lambda) = BK\log[1+\eta_P(K,\Lambda)\Lambda] \tag{31}$$

The bandwidth efficiency, denoted by $\eta_B(K,\Lambda)$, is defined implicitly as

$$TC(B,K,\Lambda) = \eta_B(K,\Lambda)BK\log[1+\frac{\Lambda}{\eta_B(K,\Lambda)}] \tag{32}$$

The power efficiency, $\eta_P(K,\Lambda)$ (bandwidth efficiency, $\eta_B(K,\Lambda)$) gives the ratio of the effective power (bandwidth) to the actual power (bandwidth) when the actual signal-to-noise ratio is Λ. The actual power (bandwidth) is the power (bandwidth) used in transmission while the effective power (bandwidth) is the corresponding power (bandwidth) needed for a single user channel to achieve the same capacity. In other words, $-10\log[\eta_P(K,\Lambda)]$ gives the power in db that we have to add to each user in order to maintain the single-user capacity. Similarly, $1/\eta_B(K,\Lambda)$ gives the ratio that we have to increase the bandwidth in order to maintain the same information rate.

Theorem 3.1.

The power efficiency satisfies,

$$\lim_{\Lambda\to\infty}\eta_P(K,\Lambda) = 0 \tag{33}$$

A lower bound to the bandwidth efficiency, $\eta_B(K,\Lambda)$, for all signal-to-noise ratio under the equal-power constraint is

$$\eta_B(K,\Lambda) \geq \frac{1}{\sqrt{f_K}} \tag{34}$$

where f_K is defined in (8).

Proof.

From the definition of $\eta_P(K,\Lambda)$, we have,

$$\prod_{n=1}^{N}(1+h_n(\lambda)) = [1+\eta_P(K,\Lambda)\Lambda]^{K\gamma} \tag{35}$$

where N, $h_n(\lambda)$ and γ are all optimally selected for that Λ. Subtracting (14) from (15), it is easy to get

$$h_n(\lambda) \leq \gamma K\Lambda(\gamma^2-1) \qquad n=1,\ldots,N. \tag{36}$$

Substituting (36) into (35), and dividing both side by $\Lambda^{K\gamma}$, we have, in the limit as $\Lambda\to\infty$,

$$\lim_{\Lambda\to\infty}\Lambda^{N-K\gamma}\prod_{n=1}^{N}\left\{\frac{1}{\Lambda}+\gamma K(\gamma^2-1)\right\} \geq \eta_P(K,\Lambda)^{K\gamma} \tag{37}$$

If $\gamma \to 1$ as $\Lambda \to \infty$, the second factor on the left hand side of (37) tends to 0, while the first factor tends either to 0 ($N < K$) or to 1 ($N = K$). On the other hand, if $\gamma \to \alpha > 1$ as $\Lambda \to \infty$, the first factor on the left hand side of (37) tends to 0 for any $N \leq K$, while the second factor is bounded. Therefore, in both cases, the left hand side tends to 0 and since $1 \leq \gamma \leq \sqrt{f_K}$, we have $\eta_P(K, \Lambda) \to 0$ as $\Lambda \to \infty$.

Substituting $\gamma = \sqrt{f_K}$ in Theorem 2.2, we have

$$\mathrm{TC}_{\mathrm{EP}}(B, K, \Lambda) \geq \frac{1}{\sqrt{f_K}} BK \log[1 + \sqrt{f_K}\Lambda] \tag{38}$$

Since the right hand side of (32) is monotonically increasing in $\eta_B(K, \Lambda)$, we have (34) when compared to (38). ∎

The TC is obtained by optimizing the balance between the "symbol rate" factor, B/γ, and the "information sent per symbol" factor, $\log[\cdots]$. As the average signal-to-noise ratio tends to infinity, the "symbol rate" factor dominates and the optimal users' signature waveforms are asymptotically identical. Then, the product term of the signal-to-noise ratios inside the log function in the TC becomes relatively small, and the asymptotic power efficiency is equal to zero. The bandwidth efficiency indicates the increase in bandwidth needed to maintain the same user rate when a single-user channel is shared by K users.

In Figure 3 and 4, we plot the Power and Bandwidth efficiency for different values of K and Λ. Also, in the same graphs, we show the lower bound to the efficiencies for the equal-power constrained channel. It shows that regardless of the signal-to-noise ratios, increasing the bandwidth by a factor of 10, we can accommodate about 50 users on the multi-user channel. This indicates a 80 percent reduction in the bandwidth required by Frequency Division Multiple Access (FDMA). Clearly, the tradeoff is a more complicated demodulating and decoding process in the Synchronous Code Division Multiple Access (CDMA) channel, which is a special case of the current model.

REFERENCES

[1] Verdú, S., "Capacity region of Gaussian CDMA channels: the symbol-synchronous case," *Proc. of the Twenty-fourth Allerton Conference on Communication, Control and Computing*, Allerton, IL, pp. 1025-1034, October, 1986.

[2] Verdú, S., "The capacity region of the symbol-asynchronous Gaussian multiple-access channel," *IEEE Transactions on Information Theory*, vol. IT-35, pp.733-751, July, 1989.

[3] Amoroso, F., "The bandwidth of digital data signals," *IEEE Communication Magazine*, vol. 18, pp. 13-24, November 1980.

[4] · ›or, D., "Theory of communication," *J. IEE*, vol. 93, pt. 3, pp. 429-457, November 1946.

[5] Nuttall, A. H., "Minimum rms bandwidth of M time-limited signals with specified code or correlation matrix," *IEEE Transactions on Information Theory*, vol. IT-14, pp. 699-707, September 1968.

[6] Narayan, P. and D. Snyder, "Signal set design for bandlimited memoryless multiple-access channels with soft decision demodulation," *IEEE Transactions on Information Theory*, vol. IT-33, pp. 539-556, July, 1987.

[7] Cheng, R. S. and S. Verdú, "Optimal Signal Design for Band-Limited PAM Synchronous Multiple-Access Channels," *Proc. of the Twenty-third Annual Conference on Information Sciences and Systems*, Baltimore, ML, pp. 321-326, March, 1989.

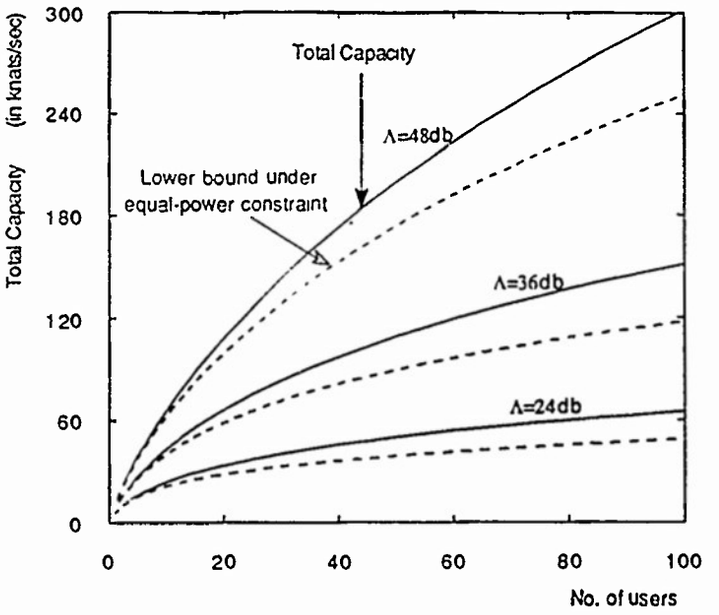

Figure 1. Total Capacity for different number of users and average signal-to-noise ratios when B=1khz

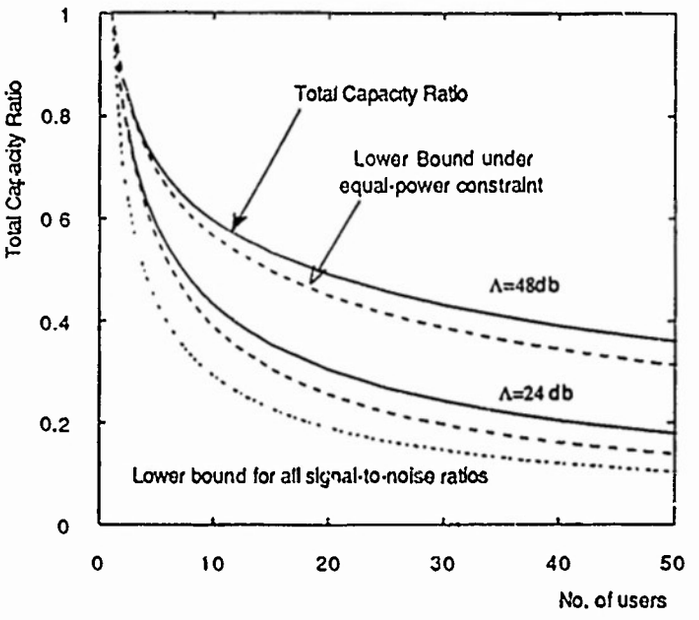

Figure 2. Total Capacity Ratio for different number of users and average signal-to-noise ratios

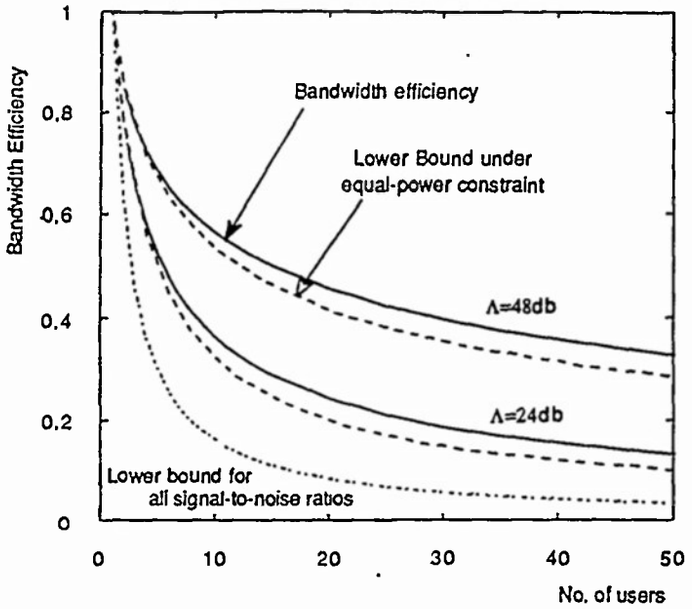

Figure 3. Bandwidth efficiency for different number of users and average signal-to-noise ratios

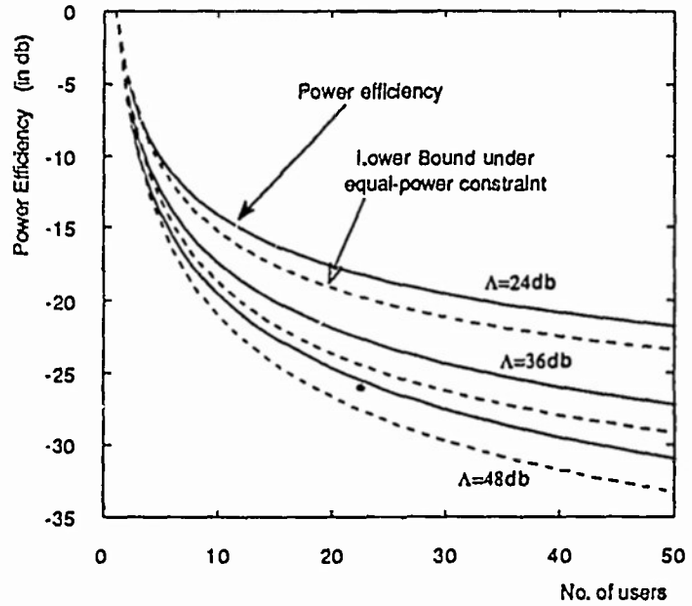

Figure 4. Power efficiency for different number of users and average signal-to-noise ratios